nature

www.nature.com/nature

Vol 449 | Issue no. 7162 | 4 October 2007

Time for a fresh start

Fifty years after the momentous launch of Sputnik, Russia has yet to find a scientific system that is anything close to meeting its twenty-first century needs.

The launch of the Sputnik probe on board an R-7 rocket on Friday 4 October 1957 can be seen, in retrospect, as the former Soviet Union's finest hour. It was the point at which that nation's alleged devotion to science and technology found its ultimate symbolic triumph.

But the awe that the launch inspired around the world (see page 538) proved to be short-lived. The country's centralized system for managing science and technology was subsequently impaled, as the Marxists would put it, on its own contradictions.

Sixteen years after the Soviet Union ceased to exist, science in Russia has yet to fully discard the baggage it inherited from the Soviet era (see pages 524 and 528). If anything, after several faltering attempts at reform, the Russian Academy of Sciences is retreating into its sclerotic past.

"Neglect of reform will cement only the decline of the Russian Academy of Sciences into mediocrity, and foster yet more emigration of young talent."

As recently as the 1980s, according to most estimates, the Soviet Union had more scientists and engineers than the United States. Isolated as they were from contemporaries in much of the rest of the world, these researchers clocked up a number of noteworthy achievements, in spheres ranging from rocket technology to pure mathematics. But even at the time Sputnik took off, the writing was on the wall for Soviet scientific advancement: the politically inspired aversion to the study of genetics, implemented by Trofim Lysenko, director of Soviet biology under Joseph Stalin, was in itself sufficient to ensure that many of the major scientific advances of the second half of the twentieth century would take place with little or no Soviet involvement.

Indeed, the success of Sputnik arguably accelerated the relative decline of Soviet science and technology by creating a massive impetus for the state sponsorship of both in the United States and elsewhere (see page 542).

After the fall

But the scientific community within the Soviet Union retained an influential and relatively comfortable position until the union's collapse in 1991. Given the speed at which events unfolded, it is hardly surprising that scientists were unprepared for the sudden end of a political system that they had occasionally criticized, but that had fed them well. Abandoned by the government, and unable to come to terms with the reforming frenzy in the early days of capitalism, Russia's scientific élite, in particular, saw the end of communism as a kind of existential threat. But it is disappointing and surprising that even now resentment, bitterness and attempts to defend old privileges are still widespread among Russian professors.

The same is true, on an institutional level, for the Russian Academy of Sciences — now, as then, the backbone of the country's research system. The academy's measurable scientific output is in decline, and yet many (perhaps most) of its members are stubbornly opposing all proposals for reform, whether they emanate from the academy itself or from Vladimir Putin's government.

The academy's autonomy is threatened by a 2006 decree from Putin that would give him the power of veto over its elected presidency. But to protect itself effectively, the academy needs to reinvent itself as a strong, modern institution, whose values are worth defending (see page 536). Clinging to past glories just won't do. The Russian government is increasingly unwilling to tolerate the academy's never-ending dodging and manoeuvring when it comes to the introduction of performance-related funding mechanisms, more competition and increased accountability.

The catastrophic funding situation of a decade ago, when many institutes survived thanks only to foreign aid, is not the sole reason for the academy's current ills. It has neglected to adopt quality standards that all successful science nations consider vital for achieving what is commonly called 'excellence'.

Lack of perspective

Getting scientific papers published in international journals, for example, is essential for researchers from Boston to Beijing. But at most Russian institutes no one cares about publication lists and impact factors, and many scientists regard as impertinent the idea that research careers should depend on papers in 'foreign' journals. The very notion of peer review is still not accepted by the old guard.

All this must change. The academy needs to make sure that the limited funds are being allocated, in a competitive and transparent way, to the best groups and projects in each field. This is even more important now that public research spending is increasing, thanks to Russia's booming economy. But in the absence of real competition and stringent quality-management the extra money is likely to dissolve in nepotism and ill-reviewed projects.

A thorough evaluation of all academy institutes, by respected scientists from Russia and abroad, would be the best way to commence modernization. Further neglect of reform will cement only the academy's decline into mediocrity or obsolescence, and foster yet more emigration of young talent. The academy should instigate this process before it has a more draconian one thrust upon it by the Kremlin.

There is no need for Russia to uncritically copy 'the' Western science system. There are, in fact, many such systems in operation, ranging from the single-grant system that predominates in the United States to the French approach of supporting most research under contracts between the government and a network of permanently staffed research institutions. Given its own scientific tradition and circumstances, the latter may suit Russia better. But the failure to modernize thus far suggests that it will be many years before Russian science and technology again holds the world in awe. ■

©2007 Nature Publishing Group

On the paper trail

The National Science Foundation's efforts to audit time-keeping could serve a useful purpose.

Griping about the paperwork involved in grant proposals is a popular pastime in academia. And there will doubtless be plenty more complaints in the United States as the National Science Foundation (NSF) launches an audit of how accurately researchers at 30 universities report the time spent on their research projects — a practice known as 'effort reporting' (see page 512).

US scientists already do plenty of bookkeeping on their federal grants, and the NSF's more rigorous requirements for them to account for their time are unlikely to be well received. Science, after all, thrives on researchers being free to explore fresh avenues of discovery without being hampered by too much bureaucracy. A careful balance must be struck between agencies and the scientists to whom they provide grant money, based on a degree of trust that funds will be used honestly and for the purposes described in the grant application.

Yet the NSF, with a budget of $6 billion, is now a major agency with a responsibility to the public to account for how that money is spent. So scientists should cooperate generously with the new audits, even if they cannot learn to love them.

Relatively little is known about how researchers spend their time using grant money. The NSF is trying to ensure that universities are supervising their researchers sufficiently and are thus avoiding flagrant abuse of the funding system. Two initial audits have uncovered sloppy record-keeping that could hint at a bigger problem. Auditors have found, for instance, that researchers say they expect to devote a certain percentage of their time on a particular grant-funded project, but then fail to keep track of the hours they actually do spend on it.

The NSF hasn't decided to conduct the audits arbitrarily. Whistleblowers at two universities highlighted cases in which researchers failed to tell the National Institutes of Health (NIH) about the amount of time they spent on projects it had funded. The universities involved subsequently repaid the money.

In the coming weeks, the NSF will finalize and release audits of five more major universities: these may bring to light further problems in the time-reporting system. As scientists complain about tight budgets and more competition for award money, they should support efforts to make sure that money is tracked accurately.

Faced with the paperwork involved in each grant application, and unsure of the details of how each project will actually pan out, researchers are often inclined to leave blank the line that requests an estimate of how much time they will spend on the project. But they should make every effort to provide this figure. It could, for instance, cut down on the number of unfortunate instances in which researchers over-commit to various proposals, and then fail to spend enough time on any of them. Eliminating this kind of overload should also help ensure that graduate students and postdocs get adequate guidance from their supervisors.

"Scientists should support efforts to make sure that money is tracked accurately."

According to NSF auditors, grant reviewers are likely to start paying more attention to effort reporting on applications within the next few grant cycles. This would be a welcome development. Congress should give the NSF inspector-general's office the money it says it needs to hire more auditors. The NSF should not, of course, go overboard in persecuting competent researchers for minor violations of bureaucratic standards. But it should try to raise the level of compliance in effort reporting.

These NSF efforts may, in time, serve as an instructive template for the NIH, a far larger agency that has done little as yet to monitor its grantees' effort reporting. With so much money at stake, a little extra paperwork is not necessarily a bad idea. ■

Living with the heat

Science and politics need to engage more than ever.

Whether face-to-face with a president or prime minister, or participating in public consultations, scientists interacting with politics must cope with two major challenges. They are likely to encounter a degree of unfamiliarity with science and the way it works that requires all their skills to convey crisply what they know in a way best attuned to their listeners. At the same time they will soon realize that those with whom they are engaging are not only in the driving seat but often have constraints, values and goals that the scientists need to understand and, where appropriate, to which they must adapt.

Throughout history, individual scientists have found themselves in direct dialogue with politicians, and have encountered the full range of acceptance, hostility, knowledge deficits, political judgement and dilemmas. Some find themselves in discussion with other stakeholders in assessing policy issues. The recent UK nanotechnology dialogues (see *Nature* **448,** 1–2; 2007) illustrated the virtues of engagement in a number of local contexts. They also showed how much both scientists and others can learn from such engagements, whether about the realities of African villages, public values or what science itself may be able to contribute.

Whether in planning future agriculture (see page 518) or in responding to potential disasters, scientists find themselves increasingly involved in political processes. This is as it should be. And not all scientists are 'natural'. Social scientists can also bring a great deal to the table, in analysing people's attitudes and behaviours.

Nature will fully explore these sorts of engagements in a series of essays, 'Science and politics', launched this week by Richard Garwin (see page 543). His account of the decline in mechanisms for providing the US government with scientific advice is pessimistic, if all too timely. More upbeat is next week's dispatch from Hans Wigzell about the changes he wrought to stem-cell policy-making in Sweden by enthusing ministers about research. And we promise that the series will end with a sense of confident affirmation of the virtues of rationality. Mindful of President Harry S. Truman's advice, scientists need to stay in the kitchen, recognizing how hot it can get. ■

©2007 Nature Publishing Group

RESEARCH HIGHLIGHTS

Volcanic paintings

Atmos. Chem. Phys. **7,** 4027–4042 (2007)

The ash from large volcanic eruptions can cool Earth's climate and cause vivid red sunsets. But the historical effects of volcanic activity on climate are difficult to quantify owing to a lack of direct atmospheric observation.

Christos Zerefos at the Academy of Athens in Greece and his colleagues have come up with a novel way to estimate the volcanic particle content in middle latitudes of the Northern Hemisphere during the period 1500–1900. They analysed sunsets captured in paintings by artists such as John Mallord William Turner, Edgar Degas and Gustav Klimt. (Pictured is one of their subjects, *The Stages of Life* (1835) by Caspar David Friedrich.)

The researchers calculated red/green colour ratios from digital photos of hundreds of paintings. They found that artists used more red after major volcanic eruptions such as those of Krakatau in 1680 and 1883. The team then used the red/green ratios to reconstruct how much light was intercepted by volcanic ash in the air. The resulting time series correlates well with and could complement existing historical reconstructions of atmospheric composition.

BRIDGEMAN ART LIBRARY

ASTRONOMY

Bright start

Astrophys. J. Lett. **667,** L37 (2007)

Not all supernovae are created equal, according to an analysis of 169 stellar explosions.

Andrew Howell and his colleagues at the University of Toronto in Canada analysed the peak brightness of one type of supernova at different times during the Universe's history. Type 1a are thought to be of almost uniform brightness, and have thus been used to probe the distances between objects in space and the strength over time of dark energy — the force pushing the Universe apart.

However, Howell found that, on average, Type 1a supernovae from the young Universe were brighter for longer than their counterparts in the current era. Correcting for this does not significantly change current measurements of dark energy, but the effect may limit supernovae's usefulness as a gauge of dark energy in the past, the authors report.

GENOMICS

Weak to strong

Genome Res. doi:10.1101/gr.6395807 (2007)

It seems that evolution may indeed favour the strong over the weak, according to Katherine Pollard at the University of California, Davis, David Haussler at the University of California, Santa Cruz, and their colleagues.

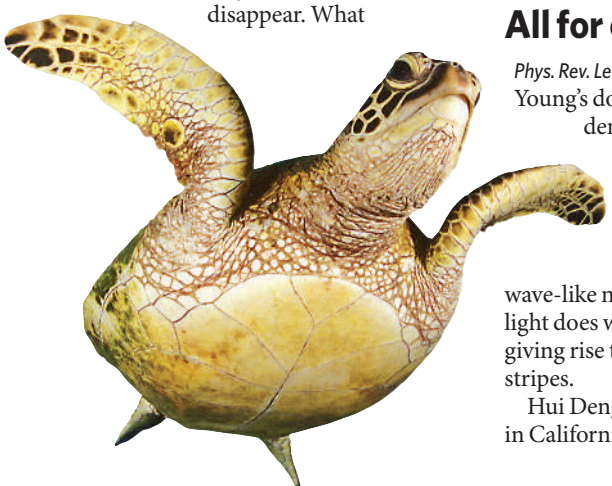
Last year, the team reported that the 200 or so most rapidly evolving spots in the human genome had more single-base changes from A and T to G and C bases — known as 'weak-to-strong' substitutions. Genome wide, mutations happen at approximately equal rates in both directions. Now, Pollard and Haussler extend their analysis and find the weak-to-strong substitution bias exists in all the many thousands of regions of the human genome in which evolutionarily recent substitutions are densely clustered. The most biased clusters contain a disproportionately high number of genes. The authors also find an intriguing correlation with areas that have high recombination rates in males.

POPULATION ECOLOGY

Wayward youth

Biol. Lett. doi:10.1098/rsbl.2007.0394 (2007)

After emerging from their nests, green turtle hatchlings bolt for the sea and disappear. What they do in the 'lost years' that follow, before they reappear to feed on seagrasses and algae near the shore, has been revealed by analysis of scute, the fingernail-like coating on turtle shells.



M. USHIODA/IMAGEQUESTMARINE.COM

Kimberly Reich and her colleagues at the Archie Carr Center for Sea Turtle Research, University of Florida, Gainesville, measured isotopes in layers of scute laid down as turtles' shells grow. High levels of nitrogen-15 in the oldest scute suggest that the turtles spend their first three to five years as carnivores, because this isotope accumulates towards the top of the food chain. Older scute also contains low levels of carbon-13, indicating that the turtles had lived in open ocean habitats. Shallow water is relatively rich in carbon-13, because the rate of photosynthesis, which preferentially snaps up carbon-12, is high.

QUANTUM PHYSICS

All for one and one for all

Phys. Rev. Lett. **99,** 126403 (2007)

Young's double-slit experiment — a classic demonstration of the wave nature of light — has been repeated for a 'polariton condensate' to test claims about such condensates' quantum nature.

Objects that have a quantum wave-like nature are expected to behave as light does when it passes through two slits, giving rise to a pattern of dark and bright stripes.

Hui Deng, then at Stanford University in California, and her colleagues looked

©2007 Nature Publishing Group

for this effect in polaritons, which each comprise a photon and a bound pair of charges, in a semiconductor device at low temperature. They confirmed that widely spaced polaritons share an almost pure quantum state. Atoms in the ever-intriguing Bose–Einstein condensates share one pure quantum state, but debate is likely to continue over whether polariton condensates deserve the Bose–Einstein label, says Deng.

NEUROBIOLOGY

Sensory deprivation

Nature Neurosci. doi:10.1038/nn1978 (2007)

Visual deprivation during the first five months of life may permanently damage interactions between the body's audio and visual systems.

Lisa Putzar of the University of Hamburg in Germany and her colleagues examined people born with dense cataracts who had no pattern vision for at least five months, but then gained good vision. Compared with controls, those born with cataracts were less likely to be distracted by a burst of noise interrupting a test containing visual cues — a sign that interference between the audio and visual systems is reduced. Those who had cataracts at birth also showed poorer audio–visual integration: they were less able than controls to recognize the words of speakers in videos with degraded audio tracks despite having similar lip-reading skills.

ARCHAEOLOGY

Keep paddling

Science **317,** 1907–1911 (2007)

The first inhabitants of Hawaii are thought to have originiated from Polynesia. The discovery of an adze — an axe-like tool — made from Hawaiian basalt on a Tuamotu atoll in East Polynesia provides the first material evidence that ancient voyagers made a round trip of more than 8,000 kilometres from the South Pacific to Hawaii and back again. Scientists long thought that journeys between remote Pacific islands must have been accidental or one-time events, although oral histories and recent research hinted otherwise.

Kenneth Collerson and Marshall Weisler at the University of Queensland in Australia examined trace elements and isotope ratios in stone adzes found on the Tuamotus islands. The adzes were from a number of Pacific islands, with one bearing the chemical fingerprint of Hawaii. Collerson suggests that pieces of rock may have been taken from island to island as a memento by travellers commemorating their long and arduous journeys.

CLIMATE MONITORING

Melt maps

Geophys. Res. Lett. **34,** L18504 (2007)

Microwave radiation emitted by snow and ice can indicate whether the surface is dry or wet. Wet surfaces imply melting. In an analysis of satellite data collected in Antarctica between 1987 and 2006, researchers confirm previous reports of extensive inland melting in early 2005 and document complex patterns of coastal and inland melting over the years. A team led by Marco Tedesco, of the University of Maryland in Baltimore, looked only at periods of 'persistent melting', defined as a period of one consecutive day and night or three days in a row. The reseachers used a new method that takes advantage of a particular microwave frequency. Tedesco says he used the same method to determine that high-altitude melting in Greenland set a record in 2007.

G. CHAPELLE/AFP/GETTY

MASS SPECTROMETRY

Hands off

Angew. Chem. Int. Edn doi:10.1002/anie.200702200 (2007)

A new technique allows the chemical make-up of human skin, plant tissue, frozen meat and other living materials to be determined quickly without harming the biological surface under scrutiny or having to treat it before analysis.

Renato Zenobi and Huanwen Chen at ETH Zurich in Switzerland and their colleagues blow a stream of nitrogen, a neutral gas, across a surface and collect a sample of what is dislodged. The sample is then transported to an extractive electrospray ionization mass spectrometer, where it can be directly analysed for its chemical composition. Using the technique, it is possible to quickly detect, from a human hand, the difference in caffeine levels before and after a person drinks a cup of coffee.

JOURNAL CLUB

Manyuan Long
University of Chigaco, Illinois, USA

An evolutionary geneticist is surprised by genes of unknown origin.

I once thought that, like us, every gene must have a mother. But recent work has identified some genes that seem to have no genetic ancestry. These 'motherless' genes pose a new challenge to understanding the molecular mechanisms and evolutionary forces that shape our DNA. This isn't the first time we've had to revise our ideas about gene evolution.

About 40 years ago, geneticist Susumu Ohno proposed that new genes originate when an existing gene duplicates, then one of the copies evolves a new function. Working with Chuck Langley in the early 1990s, I had the luck to discover a gene in flies that added another strand to Ohno's story. The gene, named *Jingwei*, is a chimaera that formed through the combination of two existing genes.

Since then, researchers have identified many other 'new' genes assembled from unrelated genes and mobile DNA elements. Often the sequences' origins can be identified. When they can't, researchers have simply assumed that subsequent evolution has masked the relationship of the gene to its ancestral sequences.

But this is unlikely to be the case for *hydra*, a gene found recently in *Drosophila melanogaster* and closely related species (S.-T. Chen *et al. PLoS Genet.* **3,** e107; 2007). No homologous sequences are found in a species that diverged from those carrying *hydra* only 13 million years ago — too recently for mutations to have obscured any related sequences. This implies that *hydra* arose *de novo*.

Another group has found a further 16 *de novo* genes in flies, which they propose evolved from non-coding DNA (D. J. Begun *et al. Genetics* **176,** 1131–1137; 2007 and M. T. Levine *et al. Proc. Natl Acad. Sci. USA* **103,** 9935–9939; 2006). These genes beg further study: what initiated their formation?

Discuss these papers at **http://blogs.nature.com/nature/journalclub**

©2007 Nature Publishing Group

habitat is probably present.

Buford Price and Robert Rohde, at the University of California, Berkeley, may have identified this missing habitat. They calculated that enough molecules such as carbon dioxide, oxygen, nitrogen and methane can diffuse through ice to sustain life.

By scanning ice cores with laser fluorimeters they detected protein spikes, some of which were indicative of single isolated cells, in just such habitats.

ASTROCHEMISTRY

Salty stars

Astrophys. J. **668,** L131–L134 (2007)

Researchers in the United States have found a dash of the unexpected in oxygen-rich stars. Lucy Ziurys and her colleagues at the University of Arizona in Tucson used the Submillimeter Telescope on Mount Graham and the 12 Meter Telescope on Kitt Peak, both operated by the Arizona Radio Observatory, to observe two red-giant stars that have shells dominated by oxygen. By analysing the recorded spectra, the team determined that the shells contain NaCl, which has previously been observed only in carbon-rich red giants.

The findings suggest that oxygen-rich stars, like their carbon-rich cousins, may be home to the complex types of chemistry that create molecular precursors to life.

BIOCHEMISTRY

Keeping the 'code'

Cell **131,** 58–69 (2007)

Certain chemical changes, or marks, made to the histone proteins around which DNA wraps seem to tell the cell whether or not that DNA should be transcribed.

Teams led by Matthias Mann at the Max Planck Institute for Biochemistry in Martinsried, Germany, and Marc Timmers at the University Medical Centre Utrecht in the Netherlands looked for proteins that bind to one chemical mark — trimethylation of lysine 4 on the histone H3. This mark is usually associated with transcriptional activity, and they found that a component of the transcription factor TFIID bound it tightly.

Dimethylation of a nearby arginine residue inhibited this binding, and other specific marks strengthened it, lending credence to the hypothesis that a combinatorial 'histone code' determines how cells read their DNA.

PLANT ECOLOGY

Grass attack

J. Ecol. doi:10.1111/j.1365-2745.2007.01307.x (2007)

Looking for signs of biological warfare past, Carolyn Malmstrom of Michigan State University in East Lansing and her colleagues delved into herbarium specimens at two University of California sites and extracted some of the oldest plant-virus RNA ever recovered.

Although ecological theory generally says that invasive species are successful outside their home ranges because they are freed from the pathogens that evolved to plague them, Malmstrom and colleagues suspect that a historical takeover of California grasslands by Eurasian grasses succeeded in part because the invaders brought viruses with them that affected the natives or changed the dynamics of an existing virus population.

They extracted barley yellow dwarf virus RNA from several specimens, including a 1917 invasive wild oat, proving that the virus was present at the time of invasion.

VISION

A scaffold in new light

EYE OF SCIENCE/SPL

Cell **131,** 80–92 (2007)

The fruitfly protein INAD had long been considered to be a scaffolding protein, organizing important visual signalling proteins that attach to it. But recent research suggests that INAD directly regulates visual perception.

Rama Ranganathan, of the University of Texas Southwestern Medical Center in Dallas, and colleagues show that, in response to light, one of five structural 'PDZ' domains of INAD transiently switches from a reduced to an oxidized state, distorting INAD's ability to bind to other molecules. This seems crucial to visually mediated reflex behaviours and for terminating visual responses.

Many scaffolding proteins contain PDZ domains, which could undergo similar conformational changes to that of INAD. Thus, rather than support components, these might serve as control centres for other signalling molecules.

Correction

The Research Highlight 'Volcanic paintings' (*Nature* **449,** 510; 2007) wrongly named Joseph Mallord William Turner as John Mallord William Turner.

JOURNAL CLUB

Andre Geim
University of Manchester, UK

Imploding atoms have softened this experimentalist's teasing views on theoretical physics.

As an experimentalist, I instinctively dislike theory papers. Too many of them seem to be written for the sole purpose of showing off an integral larger than a competitor's, or to present multiple theories just in case one idea proves right and so is hailed as visionary. I feel even less warmly towards theories that are nigh on impossible to check, such as the supposed precursor to a theory of everything, string theory.

But speaking seriously, even the most obscure predictions can turn out to be spectacularly relevant.

In our lab we have been studying graphene, a material that comprises a single layer of carbon atoms arranged similarly to chicken wire. Because electrons in this material mimic ultra-relativistic particles, it should be possible to observe in their behaviour century-long-predicted phenomena such as the Klein paradox (which concerns how highly energetic electrons tunnel through supposedly impenetrable barriers) and *zitterbewegung* (jittery movements of relativistic wave-packets).

Several recent theory papers on the physics preprint server arXiv predict another coup for graphene (see A. V. Shytov *et al.* arXiv:0708.0837; 2007).

According to relativistic quantum theory, atoms containing more than 170 protons cannot exist, because electrons around nuclei with such a large charge would fall into the centre. Nuclear physicists have not come close to creating atoms heavy enough to test this prediction. But the recent theory papers suggest that it should be relatively easy to observe the effect in graphene. This is because electrons in this material interact much more strongly than they do in atoms, so should fall down on charged impurities (standing in for nuclei) rather routinely.

This makes me wonder: could we design condensed-matter systems to test the supposedly non-testable predictions of string theory too?

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

©2007 Nature Publishing Group

NEWS

Researchers criticized for poor time-keeping

With billions of dollars flowing into US universities, monitoring how researchers spend their time is itself almost a science. Today's active researcher, who teaches, consults and works far beyond a 40-hour week, can be a hard animal to track.

Now, the US National Science Foundation, based in Arlington, Virginia, has launched the most comprehensive inquiry yet into how researchers account for their time on government grants. Concerned about how universities oversee scientists' hours, the agency's Office of Inspector General is in the first phase of a nationwide audit of 30 major universities.

Every year, about one-third of award funds from the National Science Foundation goes to salary and wages, for a total of around US$1.3 billion. "This audit is designed to get the big picture to see where the money goes," says Tim Cross, the foundation's deputy inspector general. "Record-keeping has been found to be good to sloppy to non-existent."

So far, two such audits have been published, for the University of Pennsylvania in Philadelphia and the California Institute of Technology (Caltech) in Pasadena. Both show deficiencies in how researchers report the use of their time and in how their institutions monitor them.

NSF

"Record-keeping has been found to be good to sloppy to non-existent."
— Tim Cross

Time is money

The auditors aren't looking to recover funds, Cross says — just making sure that universities follow the rules. Such audits can result in less money going to the institutions. For instance, cutting Caltech's indirect costs by just 0.5% saves $600,000.

Results from these first audits led the Office of Inspector General to direct private accounting firms to look into the next batch: the University of California in San Diego, the University of California in Berkeley, the University of Utah in Salt Lake City, the University of Illinois at Urbana-Champaign and Vanderbilt University in Nashville, Tennessee. All remain closed-mouth about the process. Utah and Illinois say that they have received draft audits for comment, which they declined to release. Draft audits for other universities are expected in the coming weeks.

The Office of Inspector General has also notified another group of universities to expect visits in the coming months. Agency officials declined to name the universities until the documentation has been finalized.

National Science Foundation officials decided to undertake the probes after several federal lawsuits involving projects funded by the much larger US National Institutes of Health (NIH) in Bethesda, Maryland. In two demonstrative cases, whistleblowers claimed that Northwestern University in Evanston, Illinois, and Johns Hopkins University in Baltimore, Maryland, failed to properly specify to the NIH how much time researchers had devoted to their grants.

Payback

Northwestern repaid $5.5 million in 2003, and Johns Hopkins $2.6 million in 2004, to settle the false-claims cases without acknowledging wrongdoing. The universities say that the cases were due to lapses in book-keeping. However, Robin Page West, the Baltimore-based attorney on the false-claims lawsuit against Johns Hopkins, notes that time-reporting violations can be complicated; in at least one case auditors had initially missed violations, she says.

After those lawsuits, the inspector general for the NIH conducted some audits in the northeast of the country, but didn't find much. With just 85 auditors scrutinizing tens of billions of dollars of programmes, the agency says it doesn't worry much about picking through an NIH's researcher's time. "We have to be pretty judicious expending our resources," says Jon Crowder, who directs the NIH inspector general's grant oversight division.

At least one institution — Harvard Medical School in Boston, Massachusetts — has

self-reported such a violation under NIH grants. In 2004, Harvard agreed to return $3.3 million after a researcher at an affiliated hospital improperly reported and used training funds for research.

Deep-seated

In the halls of US academia, stories abound of brilliant researchers whose time-reporting for their several big grants would add up to well over 100% of their time. Research misconduct cases have also hinted at misreporting. In one case, a scientist found to have faked data had NIH research grants with time commitments totalling much more than 100%.

In one of the few studies on reporting practices, a Minnesota sociologist has found much potential for irregularities in reporting. In a survey of nearly 1,800 NIH grantees (see *Nature* **435,** 737–738; 2005), Brian Martinson of the HealthPartners Research Foundation in Minneapolis found entrenched violations among early- and mid-career scientists.

For mid-career researchers winning their first RO1 grant, 70% said that they had used funds awarded for one project for other projects within the previous three years. For postdoctoral researchers, 30% acknowledged this practice. "This indicates this is the norm nationally," says Martinson. "This is

©2007 Nature Publishing Group

Caltech is one of the first universities to be audited by the National Science Foundation.

something the NIH may not want to know very much about."

For the National Science Foundation, the role of aggressive watchdog is new. In its first scrutiny, at Pennsylvania, auditors determined that the university couldn't show that at least $9.2 million, or 37%, of the salaries billed to the foundation between 2002 and 2004 hadn't actually benefited other activities instead. "The systemic nature of this control weakness raises concerns about the reasonableness and allowability of the labor effort charges on [Pennsylvania's] other $525 million of federal awards," auditors concluded. Officials at the university say that new control systems have been put in place to address the audit's findings.

In addition, in 23 of the 65 effort reports audited, managers had approved the reports without verifying that the work took as long as claimed. And 24 of the reports were filed late.

The university didn't hold department chairs responsible for the deficiencies. Who should be held responsible "is a big issue", says Theresa Ashman, controller at the University of Utah. Similar concerns have arisen, she says, in the ongoing audit at Utah, which received $21.4 million from the foundation last year.

At Caltech, grant applications showed that researchers weren't disclosing the amount of time to be spent on current and planned projects, as required by the National Science Foundation. In particular, four of five faculty members studied did not include time for the work they were proposing to do in 10 of their 11 proposals. Furthermore, some of the researchers who had committed to spending up to 20% of their time on a foundation grant failed to report this to Caltech's payroll system, which had a ripple effect of inflating the university's reimbursement rate for indirect costs. And every report checked was filed later than the 150 days allowed by Caltech policies.

Action points

Sharon Patterson, the associate vice-president for finance at Caltech, declined to speak about the specific details of the audit. Through Caltech spokeswoman Jill Perry, she stated that the university had issued a formal response to the audit at the time, which noted that Caltech's grant-management programme was sound, and that it is improving the tracking of faculty members' time-reporting.

Among the examples in the audit was that of physicist Robert McKeown, a lead researcher at Caltech's Kellogg Radiation Laboratory. McKeown's 33-person lab has received $11.9 million in grants from the Department of Energy and $3.8 million from the National Science Foundation to, among other things, develop the California High School Cosmic Ray Observatory for research and education.

According to auditors, McKeown explained that he had spent half his time on the project for the science foundation, and half on the studies for the Department of Energy. But auditors said that Caltech records could confirm only that he spent 17% of his time on research funded by the National Science Foundation. They found no records to support his time spent on the projects for the Department of Energy, on which he had pledged to spend 20% of his time. He also left blank on a proposal submission to the foundation the time that he proposed to spend on various projects. McKeown did not respond to *Nature*'s requests for comment.

"This is something the NIH may not want to know very much about."

Speed bumps

Another Caltech researcher cited for time-reporting issues is Steven Low, a computer engineer whose NetLab has led projects that have repeatedly set supercomputer speed records. Auditors say that Low told them he spent half his time on five grants from the National Science Foundation, 20% on Army and Air Force awards, and 30% on teaching and administration. But auditors found Caltech records indicating only that he spent 18% of his time on three of the foundation grants. Caltech had no records for time spent on the Air Force proposal, the Army award and the other two foundation grants. The auditors also found that none of the five grants disclosed the researcher's pledged time commitment. According to Perry, Low was too busy to comment.

Although such deficiencies could be construed as just bad paperwork, the National Science Foundation is worried that the grant-award process is being undermined. Future grant reviewers are likely to keep a sharper eye on the work commitment proposed by applicants, says Joyce Werking, an audit manager at the Office of Inspector General. "This has been below the radar," she says. "Review panels don't get into it."

Granting panels and programme managers probably will from now on. ■

Rex Dalton

See Editorial, page 508.

ALAMY

EMERGING TECHNOLOGY
Read our reports from MIT's technology conference.
http://tinyurl.com/2ufcy5

S. MCKIERNAN/NEWSCOM

©2007 Nature Publishing Group

NATURE|Vol 449|4 October 2007

F. PINI

ZOO NEWS

Sat-nav savvy
Colleagues of the late Steve Irwin (The Crocodile Hunter) were stunned when one of the saltwater crocs in their satellite-tracking study showed an amazing homing instinct — navigating more than 400 kilometres in less than 3 weeks to return to where it was captured.

SCORECARD

Ireland's smoking ban
▲ It's not just Irish people's lungs that are benefiting — the smoking ban is music to their ears, too. Instrument repairers have noticed a boost in the health of accordions, concertinas and bagpipes played in smoke-free pubs.

Greece's smoking trees
▼ Environmental group the WWF reports that vegetation will take 20 years to bounce back from the fires that raged across Greece during August, and that the fir forests will take more than a generation to recover.

NUMBER CRUNCH

460 kilograms of illegal black caviar were seized by Russian police in a raid last week.

90% is the amount that beluga sturgeon stocks have declined in the past 20 years.

US$7,500 is the price fetched by a kilogram of caviar in London, making it a tasty temptation for black-marketeers.

ON THE RECORD

"I think I could have been an art dealer. I have pretty good taste. I've always liked portraits. For me, faces dominate my life. It's not about big boobs. It's about the face. Always."

James Watson tells us what he might have done had he not turned his mind to unlocking the structure of DNA.

Sources: BMJ, Reuters, PLoS ONE, San Diego Union-Tribune

SIDELINES

A. BACCHELLA/NATUREPL.COM

Geneticists used just 21 dogs to find the genes that code for the Rhodesian ridgeback's contrary hair.

Dogs help sniff out genes

Man's best friend is becoming the geneticist's too. Researchers have made good on the dog genome's promise: a quick-and-dirty way to find the genes responsible for physical traits using just a couple of dozen pooches and a gene chip.

Kerstin Lindblad-Toh, of the Broad Institute of Harvard and Massachusetts Institute of Technology in Cambridge, Massachusetts, and her colleagues have devised a method of locating the genes responsible for specific traits that requires as few as 10 animals with the feature and 10 without — as long as they are all the same breed[1]. The team has also identified the genes that give the Rhodesian ridgeback breed its ridge but additionally predispose the dogs to a crippling developmental disease called dermoid sinus[2]. Such feats were predicted when Lindblad-Toh's team mapped the dog genome[3] but this is the first time they have been achieved.

The technique exploits the unique evolutionary history of dogs, which humans tamed from grey wolves between 15,000 and 100,000 years ago. Over centuries, humans have bred dogs, selecting for traits such as size and ability to herd sheep. Most of the 400 breeds descend from just a handful of hounds. The result: vast stretches of genetic similarity in dogs of the same breed, allowing geneticists to spot the few differences relatively easily. "Here you have the perfect genetic model," says Leif Andersson, a biologist at the Swedish University of Agricultural Sciences, in Uppsala, Sweden, who co-leads the project.

Using a gene chip that analysed about 27,000 single-letter differences across the entire dog genome, the team identified a region of 750,000 base pairs in the Rhodesian ridgeback genome that differed between ridged and unridged animals. The unridged version accounts for just 5% of ridgebacks because breeders select for the trait[1].

To identify the specific genes, the team turned to Thai ridgebacks, which exhibit the same feature without being closely related to the Rhodesians. By comparing the same DNA stretch between the breeds, the team discovered the culprit: extra copies of four genes involved in fetal development.

The finding helps explain the prevalence of dermoid sinus in the dogs. This painful development disorder, where the skin fails to separate from the nerve chord, affects nearly one-tenth of Rhodesian ridgebacks, says team-member Claire Wade, a geneticist at the Broad Institute. "Often the animals have to be put down," she says.

Dogs lacking the duplication of genes are unridged; those with one copy have a normal ridge; but having two copies also carries an 80% risk of dermoid sinus. The mechanism paves the way for geneticists to use the dog genome to help identify genes involved in disorders such as diabetes that also affect humans. ■

Ewen Callaway

P. WENGER/NATUREPL.COM

1. Karlsson, E. *et al. Nature Genet.* doi:10.1038/ng.2007.10 (30 September 2007).
2. Hillbertz, N. *et al. Nature Genet.* doi:10.1038/ng.2007.4 (30 September 2007).
3. Lindblad-Toh, K. *et al. Nature* **438,** 803–819 (2005).

©2007 Nature Publishing Group

Minimum telomere length defined for healthy cells

How long have I got left, doctor? For a cell, at least, scientists may be close to an answer. A cell's lifespan depends on the length of its telomeres — the regions of repeating DNA that protect the ends of chromosomes. Every time a cell divides, its telomeres get shorter until they become unstable and cause chromosomes to fuse together. These fusions can make the chromosomes break when cells divide, leading to cell death or triggering genomic rearrangements associated with the early stages of cancer.

Now, Duncan Baird of the University of Cardiff, UK, and his colleagues have found the shortest length telomeres can reach before they cause fusion. In addition to the steady loss associated with ageing, random mutations can cause large stochastic deletions that dramatically shrink telomeres in young, healthy cells. Baird's team focused on such cells from fetal human lung cultures whose telomeres had shrunk to the threshold length of less than 77 base pairs long. "These were normal human cells that had long, functional telomeres," says Baird. "You wouldn't expect any telomeric problems."

The critical telomere length was 12.8 repeats (of six base pairs) long, they found — any shorter and the chromosomes began to fuse together at their ends (R. Capper *et al. Genes Dev.* doi:10.1101/gad.439107; 2007). Baird suggests that without sufficiently long telomere caps, the cell may perceive the chromosome ends as broken strands of DNA and attempt to piece them together.

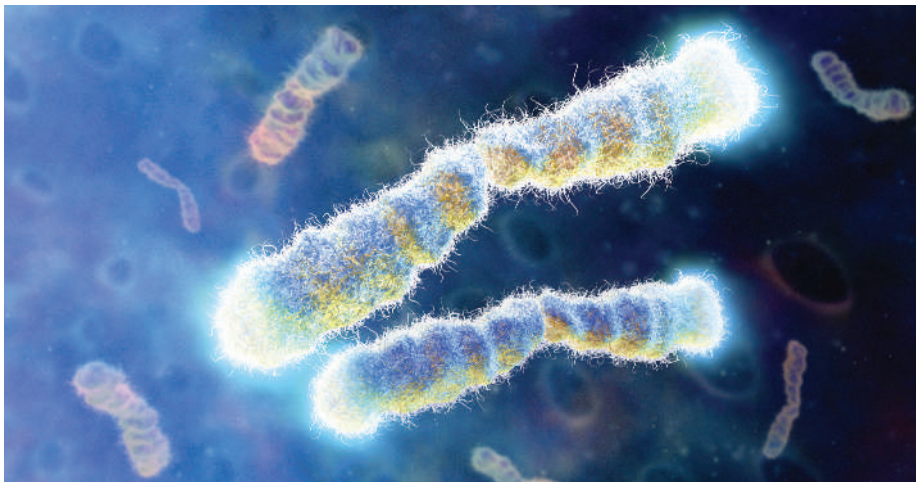
"It's important to know just how short that telomere needs to be."

The new results support the argument that it is the shortest rather than the average telomere length that counts, says Dominique Broccoli, a molecular oncologist at Memorial Health University Medical Center in Savannah, Georgia. "It's important to know just how short that telomere needs to be," she says. "It's something that people have talked about for a long time."

Although scientists suspected that this fusion might occur in otherwise healthy cells, this is the first time it has been demonstrated, says Laure Sabatier of the department of radiology and oncology at the French Atomic Energy Commission in Fontenay-aux-Roses. "This suggests that loss of large parts of telomeres would occur even during normal cell proliferation," she says. Baird estimates that telomere fusion occurs in nearly four out of every million cells in young, healthy human cells in the lab.

The researchers also gained fresh insight into the process by which chromosomes fuse across their depleted telomeres. Each end chemically stitches together in an error-prone mechanism that can itself yield large DNA deletions. The stitching mechanism — called microhomology-mediated end joining — aligns and joins the ends of the two chromosomes via tiny sections of their complementary base-pair sequences — sometimes only six nucleotides long. This represents an interesting new pathway of telomere-driven genome rearrangements, says Broccoli. ■

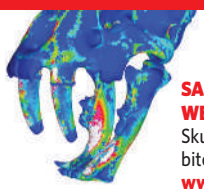
Heidi Ledford



HYBRID MEDICAL ANIMATION/SPL

Size is important: telomeres (highlighted) protect the ends of chromosomes.

©2007 Nature Publishing Group



SABRE-TOOTHED CATS WERE WEAK IN THE JAW
Skull model shows a wimpy bite for fearsome cats.
www.nature.com/news

PNAS

Water doesn't mind the gap

E. FUCHS/ GRAZ UNIV. TECH.

Researchers have managed to create a bridge of water between two beakers.

It could be the world's longest liquid bridge: researchers have coaxed water into leaping a 25-millimetre gap between two regular glass beakers in a gravity-defying stunt. The engineering feat could involve a hitherto unknown microstructure of water, the researchers say.

"Nobody would expect stable bridges to form," says Elmar Fuchs, a physical chemist at Graz University of Technology in Austria.

Fuchs's team applied up to 25,000 volts across electrodes placed in two beakers filled nearly to the brim with distilled water. Within a millisecond, water crawled up to the edge of one beaker and, in a burst of sparks, leapt across the gap between them. As the researchers moved the beakers apart, the bridge grew. The resulting thin cylinder of water stood for up to 45 minutes (E. C. Fuchs *et al. J. Phys. D* **40,** 6112–6114; 2007).

"The bridging part is new — I haven't seen that," says Hsueh-Chia Chang, a chemical engineer at the University of Notre Dame in Indiana. However, he says, the achievement is "not that surprising".

Chang uses electric fields to expel water droplets from the tips of capillaries in cone-shaped jets called electrosprays. He says that the water bridge relies on similar phenomena, on a bigger scale. The bridge takes advantage of the well-understood polarity of water molecules — due to the positively charged hydrogens and the negative oxygen. In an electric field, water molecules line up. In this case, Chang theorizes, the molecules orient themselves to form a cylindrical bridge. The repulsion between similarly charged atoms effectively wedges the bridge in place, he suggests.

Using a charged glass wand, Fuchs and his colleagues could bend the bridge from side to side. They found that water actually flowed through the centre of the bridge, from one beaker to another. The team analysed the moving water and found that it was denser than the surrounding bridge. Fuchs says he doesn't fully understand the phenomenon, but speculates that the denser water coalesces into microstructures that somehow strengthen the bridge. The group plans to carry out X-ray experiments to determine the molecular structure of the water flowing through the bridge.

The bridge lasted until too many dust particles and ions had entered the water, increasing the current across the bridge. It heated up and eventually ruptured. ■

Eric Hand

Nukes: next generation not fit for certification

The Bush administration's controversial plan to replace its cold war stockpile of nuclear weapons with a new, 'safer' nuclear arsenal has met with fresh criticism. An independent panel of scientists has concluded that "substantial work remains" before the next generation of warheads is fit for certification.

In a report released on 1 October, the JASON defence advisory group, which consults regularly for the government on a range of technical issues, said there was insufficient peer review over the design for the Reliable Replacement Warhead (RRW). The RRW is to be built without compromising a 1992 US moratorium on all nuclear tests. The report also calls for an extensive battery of non-nuclear 'subcritical' tests to ensure that the new warhead will work as required.

Panel-member Roy Schwitters, a physicist at the University of Texas in Austin, says that the design's "approach is valid", but "in lieu of new underground tests, we feel that peer review should have a bigger role" in the certification process. At present, the internal peer-review process of the three involved labs looks only at aspects of the design, not the entire system, he says.

The agency responsible for maintaining the US nuclear stockpile — the National Nuclear Security Administration (NNSA) — claims that the RRW would be more dependable than the current generation of cold war weapons. It would also contain new safeguards that would prevent its use if it fell into the hands of terrorists. But arms-control advocates say the programme is a costly handout to the nation's nuclear weapons laboratories and, they say, questions about the RRW's design could lead to renewed testing.

"Peer review should have a bigger role in the certification process."

The report is a further setback for the programme, which is already under enormous congressional scrutiny, says Jeffrey Lewis, an arms-control expert at the New America Foundation in Washington DC. "It suggests that the NNSA did a poor job in putting together a certification plan," he says.

This week, the NNSA also reported that its dismantlement of deployed warheads had accelerated by 146% over the 2007 fiscal year — three times greater than it had anticipated. Under the 2002 Strategic Offensive Reductions Treaty with Russia, the United States is committed to reducing its warheads to between 1,700 and 2,200 by 2012. ■

Geoff Brumfiel

©2007 Nature Publishing Group

France lays plans for a green future

A recipe: take a roomful of French farmers, fishermen, trade unionists, captains of industry and environmentalists, then mix in scientists and politicians. Ask them to come up with an agreed blueprint for a green revolution in France and leave them to simmer for ten weeks. It might seem a rich dish, but last week these unlikely bedfellows surprised many, including themselves, by reaching a consensus on proposals for almost everything from climate change to biodiversity. Their recommendations (see 'Key proposals') will be finalized at a summit chaired by France's president, Nicolas Sarkozy, at the end of October.

B. EDME/AP

Organic farming in France could make up 20% of production by 2020.

Environmental progress in France lags badly behind that of some of its European neighbours, such as Germany and Scandinavia. But Sarkozy made the greening of France a major plank of his election campaign this year. He has since created a superministry for ecology, biodiversity and sustainable development, with responsibility for the powerful sectors of transport, energy and construction — a first in France, where ecology was previously off the political radar (see *Nature* **447,** 518; 2007).

Scientists involved in this vast whirlwind arbitration say the exercise marks a first in the handling of science-based issues in France. Gathering the main sectors of society — including environmental groups, long shunned in France — round the same table is a new experiment in democracy, says Marion Guillou, president and chief executive of Paris-based INRA, Europe's largest agricultural research agency. The consultation is called the Environmental Grenelle, in reference to the Paris area where the 1968 government held talks with unions and industry to end weeks of rioting and strikes. It has given the greens a strong voice for the first time, environmentalists say. "It's the first time in France that we have all the main groups and the government talking ecology and the environment," says Daniel Richard, president of conservation group WWF France, and spokesman for Alliance for the Planet, a consortium of 80 French ecological organizations.

Honest brokers

"It was an historic experiment, and was a success," agrees Jean Jouzel, a climatologist who heads the Institut Pierre-Simon Laplace in Paris, and who is France's representative on the Intergovernmental Panel on Climate Change. Consensus was far from obvious in July at the start of the exercise, says Jouzel, who co-chaired the Grenelle working group on climate change with Nicholas Stern, economic adviser to the UK government and author of an influential 2006 report on climate change.

Yvon Le Maho, a biodiversity researcher at the Hubert Curien Multidisciplinary Institute in Strasbourg, felt some discomfort in such broad socio-economic negotiations. "It's unusual for scientists to be confronted so directly with the rest of the population," he explains. By the end, though, he says he was impressed by the other groups' appetite for understanding how research could help the process. Having scientists on tap, and not on top, turned out to be no bad thing either, says Le Maho. Politicians and others did a better job brokering the issues than scientists ever could, he adds.

As the weeks went by, groups found common ground, says Guillou. "There was a chemistry that worked." Everyone realized they were facing a real problem, adds Jouzel, echoing the words of Jean-Louis Borloo, minister of ecology and sustainable development: "We have no alternative but to radically change the rules and bring about an environmental revolution."

But despite the wide consensus on many areas, two major issues, the future of nuclear power — which meets nearly 80% of France's electricity needs — and the planting of genetically modified crops, eluded agreement. The meetings were too short to overcome existing entrenched positions here, Guillou says regretfully. Sarkozy has made it clear that there will be no scaling back of nuclear power. The government's final position on genetically modified crops is less clear. Although these seem set to face tighter restrictions with the proposed creation of a separate body to consider approvals on a case-by-case basis, a moratorium cannot be ruled out.

The groups' conclusions, released on 27 September, include 60 pages of recommendations, and more than 1,000 pages of conclusions from some 300 participants in 8 working groups. They are now open to public consultation on the Internet, with the government deciding on its final actions at the October summit.

But Richard knows this is only the start. In the next phase the rosy consensus among stakeholders will give way to lobbying on what gets implemented and how, he predicts. But the proposals agreed on have strengthened the hand of the sustainable-development agenda, he says. Farmers have agreed to cut pesticide use, Richard points out. "That's new; now it is a question of negotiating by how much, and by when."

Guillou agrees that the Environmental Grenelle marks a watershed in France. "Something happened," she says, "there has been a convergence." But she adds that it will require research agencies like her own to adapt. ■

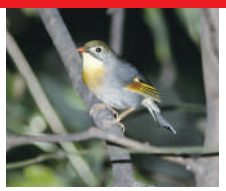
Declan Butler

Key proposals

- All newly built homes to produce more energy than they consume by 2020. Renovate all existing buildings to save energy. Ban incandescent light bulbs by 2010. Reduce greenhouse-gas emission by 20% by 2020.
- Increase renewable energy from 9% to 20–25% of total energy consumption by 2020.
- Bring transport emissions back to 1990 levels. Reduce vehicle speed limits by 10 kilometres per hour. Taxes and incentives to favour clean cars. Shift half of haulage by road to rail and water within 15 years. Develop rail and public transport.
- Reduce air pollutants quantitatively.
- Create a national network of 'green' corridors and nature reserves.
- Increase organic farming from 2% to 6% of total acreage production by 2010 and to 20% by 2020.
- Ecological groups to be stakeholders, like trade unions, in government negotiations.
- Create a body to review planting of genetically modified crops on a case-by-case basis.

D.B.

©2007 Nature Publishing Group



M. STRANGE/NHPA

ALIENS MAY BE LAST HOPE FOR HAWAII
Invasive birds are now the main reason some native forests thrive.
www.nature.com/news

A. FILATKINA/WWF

SNAPSHOT
Return of the geysers

Four months after a landslide dumped 4.5 million cubic metres of rock and mud into Russia's Valley of Geysers, two geysers have re-emerged, and others may be on their way, researchers say.

Of the 41 notable geysers in the valley, which is situated on the eastern side of the Kamchatka Peninsula, 9 were buried directly by the 3 June landslide. Another 13 were submerged when debris dammed a river. But water levels in the resulting lake have since dropped, revealing two previously submerged geysers — the dramatic 'Bolshoi' (great) geyser and the colourful 'Malachite grotto'. Meanwhile, steam is beginning to puff up around debris clogging a third geyser, says Laura Williams of conservation group the WWF.

Initial reports of the landslide worried Tamas Torok, a microbiologist at Lawrence Berkeley National Laboratory in California, who regularly visits the region to study the its unique microbial climate. "They said the Valley of Geysers was destroyed," he says. But on a recent visit, he found that many of his research sites had escaped unharmed.

Some microbiologists have adapted to the landslide by redefining their research goals, says Christopher Romanek, a biogeochemist at the University of Georgia in Athens: "Now we want to go out there and take samples to try to understand how these systems rebound after a catastrophe." ■

Heidi Ledford

Cool reaction to Bush's climate summit

H. N. GHANBARI/AP

The inaugural climate summit hosted by President George W. Bush in Washington DC last week was greeted with cautious optimism from some scientists. But his continued insistence on voluntary measures for reducing greenhouse-gas emissions met with intense scepticism from international delegates and observers. And although Bush underscored the risks posed by global warming in a public speech on 28 September, many scientists and advocates say his administration continues to oppose actions necessary to mitigate the problem.

Peter Cox, a climate modeller at the University of Exeter, UK, says he is pleased that the United States is now "raising the flag" on global warming, abandoning a previous stance that, he says, amounted to denial. "Essentially, what we had before was the uncertainties in the long-term projections, which are always going to be there, being used as a reason for inaction," Cox says. "It is a relief that people are talking about the need for action despite the fact that there are still significant uncertainties."

But in the first US-government-hosted climate summit, Bush offered no specific targets or solutions, aside from the establishment of an international fund to promote clean-energy projects in the developing world and a general endorsement of technology as the best solution. "There is a way forward that will enable us to grow our economies and protect the environment, and that's called technology," he said. The administration, which invited developing nations to the talks, also continued to push for voluntary measures that could be designed and implemented at the national level. This was seen as another effort to draw support from developing nations, whose emissions are increasing sharply, and are expected to continue to do so in the coming decades.

Bush favours a voluntary approach to lowering emissions, but others say this will be ineffective.

"You can't solve this problem on voluntary measures. It's a fraud," says climatologist Stephen Schneider, from Stanford University in California, who says that even immediate action might not prevent a rise of 2 °C above pre-industrial levels. "The science is very clear that unless we are incredibly lucky or there is an incredible breakthrough, we are going to be well over 2 °C," he adds. "The impact-assessment folks say that is when you are going to have all sorts of nasty stuff."

John Ashton, the UK climate-change envoy attending the summit, criticized the Bush administration for failing to suggest a target for the United Nations' climate negotiations that are scheduled for December in Bali, Indonesia. Many European countries are calling for a treaty that would require emissions cuts of more than 50% by 2050. "If you accept the science, if you want to be taken seriously as accepting the science, you have to be willing to say to the international community where you think the threshold of dangerous climate change lies," Ashton says. "It's very hard to reconcile that refusal to say what the goal should be with a position of leadership." ■

Jeff Tollefson

©2007 Nature Publishing Group

White House axes risk-assessment bulletin

The White House has withdrawn a controversial bulletin that would have sought to standardize the way in which US government departments use science to assess risks.

The bulletin had been fiercely criticized as an underhand attempt to weaken regulatory agencies such as the Environmental Protection Agency (see *Nature* 442, 242–243; 2006). Having been asked by the White House to review the document, the National Academy of Sciences last autumn called for its complete withdrawal.

But earlier this year, senior officials at the White House Office of Management and Budget (OMB) said that the bulletin was still in play (see *Nature* 448, 624; 2007). With senators pressing for its withdrawal, however, the OMB announced on 19 September that the bulletin is no more, and instead issued a memo updating its previous, 1995 advice on scientific risk assessment.

Lack of inspection raises concerns over drug trials

Incomplete data, lack of coordination and poor follow-up are hampering the overseeing of clinical trials by the US Food and Drug Administration (FDA), says a government report.

The 28 September report — from the inspector-general's office in the department that oversees the FDA — found that the agency has no comprehensive database of the clinical trials it requires of companies developing new drugs and devices. On the basis of information in a National Institutes of Health trials registry, the report said that the FDA probably inspected fewer than 1% of all such trial sites between 2000 and 2005.

In addition, some 75% of FDA inspections were focused on verifying the accuracy of data in previously completed trials, rather than on protecting human subjects in active trials. The FDA responded that its most important role in protecting human subjects takes place before clinical trials begin, when it reviews company protocols.

Climate sceptic quits over 'lack of academic freedom'

Patrick Michaels, one of the most controversial sceptics on global warming, retired from his post as state climatologist for Virginia earlier this summer, citing concerns about academic freedom.

Michaels, an environmental scientist at the University of Virginia in Charlottesville, which co-hosts the state's climatology office, has been criticized for playing down the risks of global warming. As a result, Virginia's Democratic governor, Timothy Kaine, last year asked him not to use the title of state climatologist.

After news of his resignation broke last week, Michaels released a statement saying it was "impossible to maintain academic freedom" because his budget — and his salary — had become politicized. Joseph Zieman, who chairs the university's environmental sciences department, counters that it was the consulting Michaels did for energy companies that got him into trouble. The University of Virginia has appointed Jerry Stenger to head the climatology office, although his official title is research coordinator.

Dawn mission to asteroid belt has lift-off at last

After surviving one cancellation and more than a year's additional delay, NASA's Dawn probe was launched into space on 27 September riding atop a Delta II rocket. The mission to the asteroid belt between Mars and Jupiter had been held up by problems including crane malfunctions and the slow arrival of instruments through the Panama Canal. It had even been temporarily cancelled owing to cost overruns.

Up and away: NASA's Dawn probe takes off. J. RAOUX/AP

If Dawn had remained stuck during October, scientists would have had to wait 15 years for another launch window, says Carol Raymond, the mission's deputy principal investigator and a planetary scientist at the Jet Propulsion Laboratory in Pasadena, California. "It's obviously a tremendous feeling to finally see this mission get off the ground," she says.

After a 2.8-billion-kilometre journey, the probe will in 2011 reach Vesta, a highly reflective asteroid. In 2015, the spacecraft will rendezvous with Ceres — recently dubbed a dwarf planet. Dawn could confirm scientists' suspicions that Ceres hosts water and a thin atmosphere.

Bird-flu virus can pass from mother to unborn child

Chinese and US scientists have found for the first time that the H5N1 avian flu virus can pass across the placenta and infect unborn children.

The autopsies of a pregnant woman and a man, reported last week (J. Gu *et al. Lancet* 370, 1137–1145; 2007), also confirm that infection with the H5N1 virus affects multiple organs, and is not restricted to the respiratory system as is seasonal human flu.

Autopsies of H5N1 victims are rare, as corpses tend to be disposed of quickly. At least three reports, however, have shown that H5N1 infection in humans can affect the brain and multiple organs, which probably explains the high pathogenicity of the virus.

See-through frog offers inside information

HO NEW/REUTERS

Who needs dissection when you can view a living frog in all its transparent glory?

Japanese biologists have created what they call the world's first see-through creature with four legs. (Some fish are naturally transparent.) The frogs started out as ordinary Japanese brown frogs (*Rana japonica*), but crossing animals with recessive genes for light-coloured skin led to transparency.

Internal organs, eggs and other normally concealed innards are all on full display in the new amphibian (pictured). Its creator, Masayuki Sumida of Hiroshima University, plans to seek a patent, according to the AFP news agency.

Correction

The News story 'Mystery ox finds its identity' (*Nature* **449,** 124; 2007) stated that Gary Galbreath and his colleagues had concluded that the kouprey was a hybrid. In fact, they revised this conclusion in March 2007, determining that the ox was a distinct species, on the basis of fossil evidence (G. J. Galbreath, J. C. Mordacq and F. H. Weiler *J. Zool.* **271,** 253–254; 2007).

©2007 Nature Publishing Group

BUSINESS

The carbon cycle ride

Carbon offset schemes in developing countries may be finding their feet after a rocky start, as **Paroma Basu** reports.

India's blooming carbon market has emerged as the hottest destination for companies in the industrialized world seeking to do their bit in the fight against global warming.

But as one of the first nations to move aggressively into carbon trading, India has also been the first to collide with the teething troubles of the fledgling market.

"There is a problem with the process of the Clean Development Mechanism (CDM) itself that no one has been able to remove," says Chandra Bhushan, associate director of the Centre for Science and Environment in New Delhi. "Independent oversight, to ensure that the processes are clean, just isn't there yet."

The CDM was established under the auspices of the United Nations Framework Convention on Climate Change. It lets nations buy certified emissions reductions, or 'carbon credits', from countries that earn them by cutting greenhouse-gas emissions.

Several issues lie at the heart of India's troubles with carbon commerce. They include the country's lack of preparedness, the United Nation's inexperience with the market, the technical difficulties of managing, supervising and regulating projects, and delays in the Kyoto Protocol coming into effect.

Despite the early setbacks, observers say that a solid system of checks and balances is now coming into place. "Because India was and currently is the CDM leader, it is natural that the problems were first faced by Indian projects," says José Domingos Miguez, an official in Brazil's Ministry of Science and Technology and a member and former chair of the CDM's executive board, which oversees the mechanism. "This is a learning-by-doing process."

> "Independent oversight, to ensure that the processes are clean, just isn't there yet."

India is home to 282 of the world's 803 registered carbon-trading projects, according to the United Nations — more than China (116) and India (108) combined. And analysts say that Indian companies have already earned about US$350 million from these ventures. With another 435 projects in the pipeline, the earnings will quickly run into billions of dollars.

Some claim that the offsetting mechanism is susceptible to easy fraud, but Indian firms counter that they have to jump through several regulatory hoops to gain registration. Under the Marrakech Accords of 2001, which fleshed out the rules for meeting Kyoto targets, projects must clear three stages of review — a process that generally takes about a year to complete.

Once a firm identifies a way to reduce emissions in its business — by switching to wind power, for example — it has to clear its idea with a government task force appointed inside each nation. Then, the projects are audited by UN-accredited 'designated operational entities', which check their methodology, documentation, finances and eligibility. Around six such entities operate in India.

In the third stage, the CDM's executive board, comprising senior climate scientists and officials from around the world, makes the final call. Projects that get the green light can finalize deals with buyers in developed nations. For the rest of the project's duration, UN-appointed auditors visit at regular intervals, awarding credits on the basis of the extent of reductions made.

Delayed take-off

Just after Marrakech, Indian companies were among the first to propose eligible projects. But the CDM board, established in 2001, didn't actually start operating properly until the Kyoto Protocol went into force, in February 2005. Even then, Miguez says, it was almost a year before the cash and people were available to monitor projects on the ground adequately.

That meant that scores of Indian firms, which filed more than 100 applications before 2005, moved headlong into the process in the absence of clear guidelines. "Everything was a confusion to begin with," recalls Subhash Rastogi, vice-president for environment, health and safety at the multi-product firm ITC in Kolkata, which moved into the carbon market in 2003 and has since registered seven projects.

Soon, the Indian carbon market began to rack up bad publicity. Questions surfaced about the legitimacy of its projects and the competence of the auditing firms appointed to oversee them. Pressure groups such as the Center for Science and Environment criticized the wrongful approval of undeserving projects.

In particular, a rule agreed at Marrakech allowed for the automatic registration of a project unless at least three board members request to review it within 4–8 weeks. "Some of the projects were going through the automatic mode and getting registered easily," says Rajesh Kumar Sethi, vice-chair of the CDM board and director of the climate-change department in India's Ministry of Environment and Forests. "The CDM secretariat was thinly staffed at the time — but those types of projects are very rare now."

Carbon credits provide incentives for facilities such as this brick factory in Lalgola, India to cut their greenhouse-gas emissions.

The secretariat, based in Bonn, Germany, now has a technical staff of about 100 to support the 20 board members, Sethi says. The board undertook three random spot-checks last year to keep tabs on national performance, Miguez adds, and it has been updating documentation requirements to tighten up the process.

Measure of doubt

An obvious problem, however, is the calculation of greenhouse-gas reductions for specific technologies. These calculations require a baseline from which to measure — but that has been tough to establish in several industry sectors, such as aluminium, cement and agriculture.

Another issue is the mechanism's crucial 'additionality' requirement, which holds that to be eligible for benefits, a project has to show that it is genuinely reducing emissions by more than it would have done otherwise. "Additionality is subjective," says Sethi. "Somebody may say it's additional while somebody else may say that it's not."

That murkiness is creating situations in

©2007 Nature Publishing Group

P. CALINESCU/PANOS

which some companies think that their projects have been rejected unfairly. "A certain randomness persists at the level of the executive board," says Suresh Iyer, deputy general manager of JSW Steel in Mumbai, which has won the world's largest credit award for any one project, for fuel-replacement at its steel mills. "There is no forum in which companies can defend their projects directly in front of the board."

"The feeling of rejection is worse when you don't know the reason," says a high-level executive at a leading cement company who didn't want to be named. "We applied for two projects with exactly the same methodology. Only one got accepted — although the rejected one is a much stronger case." Of 36 projects that the CDM board has rejected so far, 14 are from India.

"The secretariat is still not resourced adequately, either in numbers or expertise," says Paravastu Rambabu, managing director of Indian operations for the consulting firm Cantor CO2E. He adds that auditing companies face similar capacity constraints.

Furthermore, companies seeking to join the CDM pay the auditors' fees, and that's a conflict of interest, according to critics. Miguez says that such payment arrangements were laid down at Marrakech, but agrees that it would be better if the CDM board paid the auditors itself.

"With every new system, there will be pluses and minuses," notes Iyer. "But as the number of projects grows, and as companies have seen actual money coming in, awareness and confidence in the process is also growing." ■

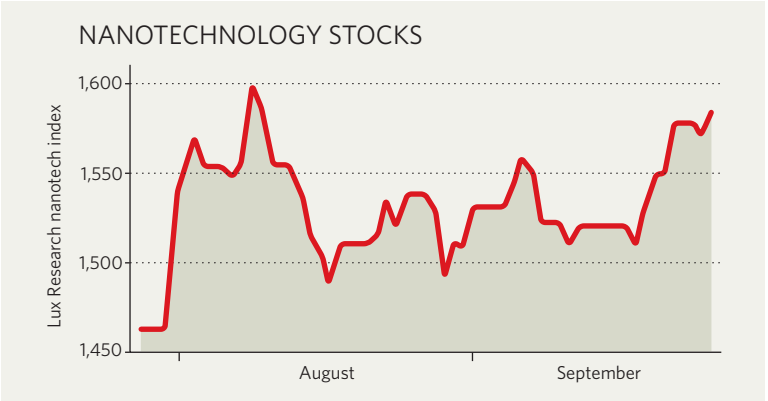
IN BRIEF

SOLAR SELL Colorado State University in Fort Collins says that a spin-off company started by one of its professors could employ up to 500 people to make solar panels that will "dramatically reduce" the price of solar power. AVA Solar will make the panels from glass coated with a film of cadmium telluride, a technology developed by Walajabad Sampath, a mechanical engineer at the university. The university claims that the panels will deliver power to the consumer for less than US$2 a watt of capacity — about half the cost of currently available alternatives.

PULL TOGETHER Leading drug companies and a group of academic institutions say that they will work together to explore people's genetic susceptibility to the side effects of drugs. The International Serious Adverse Events Consortium, led by Arthur Holder, formerly of Baxter International, is starting off with studies of susceptibility to a skin condition called Stevens–Johnson syndrome and to drug-related liver problems. Pharmaceutical companies hope that genetic tests will eventually become available that will identify patients at risk from taking certain drugs.

BIOFUELS BACKING Amyris Biotechnologies, a company in Emoryville, California, that specializes in designing microorganisms for use in the production of biofuels, has secured $70 million in funding from venture-capital firms. A tranche of backers led by Duff Ackerman and Goodrich Ventures in San Francisco, California, have come up with the money. Amyris says that the money will help it to develop fuels such as biogasoline, biodiesel and bio jet fuel, which it hopes to bring to the market by 2010.

MARKET WATCH



After the general dip that hit global stock markets at the end of July, the value of nanotechnology companies steadied in the past few months, as investors absorbed a mixed bag of financial results.

The Lux Research nanotech index, charted above, measures the performance of a basket of companies that sell nanotechnology products and services as well as some major industrial corporations that make strategic use of them.

There weren't many star performances among the companies over August and September. Peter Hebert, chief executive of Lux Research, the New York consultancy firm that compiles the index, says that Nanophase Technologies Corporation, a relatively long-established nanomaterials firm in Romeoville, Illinois, is trading well, with sales up 75% from last year. But the market remains unmoved, with its stock stuck at around $7 for the past three years.

Flamel Technologies, a drug-delivery company in Vénissieux, France, has gone through a difficult period after losing a deal with the British pharmaceutical giant GlaxoSmithKline. Flamel's stock has fallen steadily from almost $40 earlier in the year to less than $9 last week.

But according to Hebert, the sector is encouraged by a pending initial public offering on the Nasdaq for stock in Nanosphere, a spin-off company from Northwestern University in Evanston, Illinois. The molecular-diagnostics company is hoping to raise up to $100 million from the offering, the date for which is about to be set. "This is one of the first true nanotechnology start-ups to reach the market," Hebert says. "A lot of people will be looking at that — and a number of others could follow." ■

Colin Macilwain

©2007 Nature Publishing Group

THE BATTLE FOR RUSSIA'S BRAINS

The Russian Academy of Sciences has resisted pressure from czarists and communists. Can it thwart the reforms planned by Putin's government? **Quirin Schiermeier** reports.

Muscovites have a complex relationship with the bold architecture they call the "building with the golden brain". The widely visible tower near Gagarin Square in the southwest of the Russian capital, topped by a futuristic chaos of glistening copper pipes and air shafts, is the modern headquarters of the Russian Academy of Sciences, founded in 1724 by Peter the Great. For many it symbolizes the pride and glory of science in a nation that has traditionally held researchers, inventors and explorers in high esteem. But for others the overly ornate superstructure can all too easily be seen as emblematic of the arrogance of an isolated and ageing academic élite that has little to offer Russia as it moves towards a knowledge-driven modern society.

The feeling that the academy, which employs some 50,000 scientists across 418 research institutes, is in dire need of reform is widespread even among scientists affiliated with the institution. Critics say that it is controlled by an old guard of 1,250 academicians raised under a Soviet culture who cling to their privileges and rituals. The academy, they say, has become a comfortable refuge for tens of thousands of unproductive 'shadow researchers' who await their pensions while producing little or no science of any merit. In the absence of proper quality control and competition, pessimists warn, Russia will risk falling further behind up-and-coming science nations in Asia, and could even struggle to maintain its status as a second-rate science nation. The response from many leading academicians within the academy is that Western-style research with its 'publish or perish' mentality offers a rat race unworthy of the noblesse of true science.

"Ten years ago Russian scientists published ten times more papers in journals such as *Nature* and *Science* than Chinese researchers did — now the Chinese have twice as many as we do," says Alexander Sobolev, a geochemist at the academy's Vernadsky Institute of Geochemistry and Analytical Chemistry in Moscow, who is frustrated by the slow pace of reform. "The problem is that the academy just doesn't support a system in which productivity counts. Scientific careers in Russia don't depend on results, and nobody really cares about international expertise. We simply don't provide the right motivation for scientific work."

Motivation understandably waned during the

Under a cloud: the Russian Academy of Sciences is fighting proposed government intervention.

Q. SCHIERMEIER

©2007 Nature Publishing Group

1990s, after the break up of the Soviet Union, when scant scientific salaries forced talented Russian scientists to find second or third jobs outside science, or to leave research altogether. Many more emigrated to research positions abroad. During the hardest years, even top academicians officially earned little more than US$100 per month. Without generous support from foreign governments and organizations, such as the Soros Foundation in New York, set up by Hungarian billionaire George Soros, the Soviet Union's scientific heritage might have decayed beyond repair.

During this period of crumbling science budgets and galloping inflation, many were concerned about Russia's nuclear heritage and its emigrating scientists. The nation was living permanently in debt, and could hardly afford to maintain some 4,000 Soviet research organizations. As economist Boris Saltykov, the first minister of science in the post-Soviet government, said in 1992: "In Russia we have too much science." But he, like others after him, struggled to reform the academy.

Against the odds, the number of academy institutes actually rose after 1990, from 330 to 418 today. The number of academicians has also grown, at the expense of other scientific employees, and despite frequent criticisms of inertia and unproductivity. Some wonder if the academy is unreformable — a feature of the Russian landscape as enduring as the Siberian steppes.

The current science minister, Andrei Fursenko, formerly a physicist at the renowned Ioffe Physico-Technical Institute in St Petersburg, is determined to make the academy more accountable. Sources close to the ministry say that long-standing discussions over the reform of the academy are turning into a real power struggle. In January, Fursenko suggested a 'model charter', which the academy's general assembly rejected in March on the grounds that it would give the ministry bureaucratic control over the academy. The academy has proposed its own charter, but a compromise is unlikely, and so the deadlock continues. Whether the academy can resist all attempts at external and internal reform remains to be seen.

Paper trail

What is the state of Russia's scientific output today? Certainly its publication record remains bleak. "Too many people here call themselves scientists without having published a single paper in the past decade," says Boris Stern, an astrophysicist at the P. N. Lebedev Physics Institute in Moscow. In 2001, Stern received a small grant from the Russian Foundation for Basic Research in Moscow to build up a database of the scientific productivity of Russian scientists and scientific institutes. He regularly updates this 'Who's who in Russian science' with the latest information from Thomson Scientific, a citation service based in Philadelphia, Pennsylvania.

The academy's general assembly voted to oppose a new charter proposed by the government.

"The problem is that we have a totally biased system for scientific reputation in Russia," says Stern. "The formal administrative hierarchy has nothing to do with informal reputation." For example, Stern's list of the most productive institutes is topped by Lomonosov Moscow State University, which is not an academy institute. And Stern found that fewer than 50,000 Russian scientists — one-eighth of the total workforce — publish at least one paper a year. According to the 2006 World Bank Russian Economic Report, researchers in Poland, India, Brazil and South Korea generate 2–3 times as many papers per person. And patent numbers tell a similar tale: patents produced per capita are 60 times higher in South Korea than Russia.

Science minister Andrei Fursenko wants to make the academy more accountable.

G. SYSOYEV/ITAR-TASS

The low productivity has not escaped the attention of the Kremlin. As the government invests in the country's diminished science base — the overall number of researchers has declined by more than 100,000 since 1995 — expectations will rise and so focus new attention on the academy's sagging performance. President Vladimir Putin is determined to give the Russian economy, which leans towards exploitation of the country's abundant oil and gas resources, a more solid industrial base. Science will be a tool on the road to a new economy Putin said in May in a speech to the State Duma, the lower house of the Russian parliament. In July, the Duma dutifully approved a massive $7-billion investment in nanotechnology over the next five years, which Putin and his advisers think will lessen Russia's dependence on petrodollars (see *Nature* **448,** 233; 2007).

The nanotech initiative, which almost matches the $1.3 billion per year the academy receives from the government, can be seen as an attempt to develop a new state sector of science outside the academy, says Saltykov. In fear of being passed over, the academy rushed to appoint Mikhail Kovalchuk, a physicist with close ties to Putin and the head of the nano-initiative, as its 'acting' vice-president for nanotechnology. The move has been criticized as a violation of the academy's statutes, which state that only full members can be elected into leadership positions — Kovalchuk is only a corresponding member.

State control

Having already threatened the academy's pre-eminence as Russia's main basic-research organization, the government is also starting to encroach on its internal operations, which remarkably survived the communist era intact. Last December, Putin signed the new 'Law on Science and Technology State Policy' which, among other things, modified the legal status of the academy to give the government more control over its operations. The new law includes a provision that allows Putin to approve the elected president of the academy — a requirement that has dismayed some leading academicians.

Alexander Nekipelov, one of the academy's chief vice-presidents, says he thinks it was 'merely a symbolic act'. "I don't think [Putin] can actually disapprove of an academy president who has been properly elected according to our rules," he adds. Still, last November, in anticipation of the new rules, the academy postponed the election of a successor to its current president, Yuri Osipov, who has reached the age limit of 70. It's not yet clear when a new academy president will be elected — it may be as late as 2008, when the country will also choose a successor to Putin.

Nekipelov does admit to "major disagreements" with the science ministry over the future course and governance of the academy. According to Fursenko's charter, important decisions, such as creating new laboratories or closing existing ones, and on allocation of funds, would be handed over to supervisory committees, in which government representatives would hold the majority. Research money would be given to individual teams on a purely competitive basis rather than as a lump sum to the academy presidium.

©2007 Nature Publishing Group

ITAR-TASS

A coffin bearing the words 'Science of Russia' symbolizes scientists' views about funding levels.

The academy leadership, backed by most of the full and corresponding academicians who make up the general assembly, vehemently opposes the new law and subsequent charter. It has prepared its own charter, which is currently being reviewed by the government, in which the academy would remain a self-governed organization, free to allocate funds independently, and with the right to set up, modify or liquidate institutes as it sees fit.

The academy is keen that most of its budget remain as basic institutional funding — as a fixed lump sum. But it has suggested that 20% of the money it receives from the government in future will be distributed though a 'competitive' interdisciplinary programme. The ministry would like to reduce institutional funding to a minimum, and introduce performance-related funding of projects and individuals, including consideration of scientists' international publication record. Many academicians hate this idea because their low publication rate in international journals would exclude most of them from getting funding under any competitive programme. If competition is to be introduced, they would prefer the programme to be managed 'in-house', using internal expertise to evaluate grants. But the government is unlikely to accept this proposal, and a compromise is currently not in sight.

To outsiders, the academy seems to be fighting a losing battle. The ministry holds the purse strings, and can force through legislation to restructure the academy. Indeed, the government has threatened to block next year's round of promised pay rises to increase pressure on the academicians, says Nekipelov.

Average salaries for academy scientists were to jump to $1,200 per month by 2008, in exchange for the academy cutting staff by 20%. Two years ago, average monthly salaries were still as low as $200; they are now at around $500, roughly similar to that of the average industrial worker. To justify even higher salaries, the government wants to impose on the academy a performance-related system that it hopes will improve the scientific output of its workforce. Future salaries would be composed of a fixed part — the tarif — plus bonuses for measurable scientific achievements.

Nekipelov gets animated when it comes to what he calls 'paying by results'. He disagrees that a simple indicator, such as the number of publications in international journals, can measure researchers' performance, as has been proposed by the science ministry. "People would start working for 'indicators' only," he says.

"Some people think that our scientists just want to pocket salaries but not work hard, and that there is no such thing as competition in our institutes," he says. "That's wrong. The truth is that scientific work as such is a very competitive activity. All academy scientists compete for funds allocated to their institutes, while institutes compete for funds allocated to different fields of science, and so forth."

Nekipelov also disputes the widely held view that most academy researchers are isolated from, and rather indifferent about, international trends in science, and refuse to give talks or write papers in English. "We do have many unofficial relations with colleagues abroad," he says. He is also critical of the influence of international journals and foreign reviewers: "The working language in this country is indeed Russian, scientists will not start speaking English here."

Saltykov is dismayed to hear this: "If that is so, then what we are doing here has nothing to do with international science." Saltykov is in favour of radical reorganization of the academy. With the exception of a few young scientists with ties to the West, he says, the members of the academy continue to live in a Soviet-type economy and state of mind. "They say that science and the market are incompatible, whereas in fact they consider money more important than principles. And they keep demanding budget money, but would like to remain totally unaccountable."

Fresh blood

Only about one-quarter of academy scientists are truly competitive, hard-working researchers, says Saltykov. He cites an unpublished review of 1,330 academy researchers done in 2005 by sociologist Sergei Belanovsky on behalf of the Moscow-based Centre for Strategic Research, a think-tank close to the ministry of economics. The study showed that almost half of the academy's scientific workforce was unproductive dead weight. Nekipelov accepts that something needs to be done: "We do have quite a strong potential," he says, "but it's true that we need more young people, and that we need to free ourselves from some of the elders who don't do much science any more."

"People here are very experienced in the art of imitation," says Mikhail Feigel'man, a researcher at

Q. SCHIERMEIER

Outside the academy: Moscow State University is one of Russia's most productive institutes.

©2007 Nature Publishing Group

President Putin (left) appointed Mikhail Kovalchuk as head of a new nanotechnology initiative.

M. KLIMENTYEV/ITAR-TASS/REUTERS

the Landau Institute for Theoretical Physics near Moscow. "Even though open scientific competition is officially encouraged, 99% of it is fake. You write a proposal, say, and then a so-called 'competition' is created, but in which only your project will fit." The nanotech initiative, for which Kovalchuk will effectively have control over the whole budget, makes no difference he says. "It's a very feudal system in Russia, and I fear this will be impossible to change in our current situation."

However, government funding is starting to become more transparent. Its $5.3 billion federal targeted programme for 'Research and Development in Priority Areas of Russian Science and Technology 2007–12' supports mostly applied research on a competitive basis in seven fields, including the life sciences, energy and nanotechnology. The scheme is modelled after Europe's Framework Programmes and is intended to increase competitive funding for research projects.

In recent years, the academy has experimented with a similar funding model — on a smaller scale and with mixed results. The 'Scientific Programmes of the Presidium of the Academy' is a scheme to enable small teams, selected on the basis of their international publication record during the past five years, to do independent research. Sixteen fields were selected for pilot projects, but only one, the $5-million cellular and molecular-biology project, has had much success.

When they were first set up in 2002, the idea of competition was almost a revolution in the academy system, and it was attacked from many sides. Only a few 'good' people applied in the first round, and most did not believe that there would be honest competition and no corruption, says Georgii Georgiev, former director of the academy's Institute of Gene Biology in Moscow and project coordinator for the cellular and molecular biology scheme.

But news got around that the programme could work, and that the grants — of $150,000 per year — could provide talented scientists with a level of scientific independence and flexibility formerly unknown in Russia. About 100 life scientists, of which 10 have since returned to Russia from abroad, are being funded in the cellular and molecular biology programme, and the second round, from 2004–08, was oversubscribed five-fold. "We improved the system by and by, so that we are considered an example as to how competition can help make our science better and more transparent," says Georgiev. The third round of the competition will start in December.

Isolated innovators

Sobolev, a corresponding member of the academy who has close ties with the Max Planck Institute of Chemistry in Mainz, Germany, has lobbied hard in the past few years to set up and expand competitive programmes to support international-level research groups in Russia. But he got little support for his ideas from decision-makers within the academy.

> **"We have a totally biased system for scientific reputation in Russia."**
> — Boris Stern

The problem is that only a small fraction of Russian scientists — those well integrated in the international community and who regularly go to international meetings — really want to change the system, explains Evgeny Antipov, a chemist at Lomonosov Moscow State University. "By and large, impact factors and publications have zero impact here on grants and careers," he says.

Konstantin Severinov, a young molecular biologist at the Institute of Molecular Genetics in Moscow who is also a tenured professor at Rutgers University in New Jersey, is one such scientist who has recently returned to Russia. He would like to transform the system from within — with the help of the next generation of Russian scientists.

"The students here are just wonderful," he says, but he points to a worrying increase in state interference in scientists' lives (see *Nature* **449,** 122–123; 2007). Scientists now have to report all contacts with foreign visitors to their institutes, for instance. Severinov was himself interviewed by the Federal Service Bureau in May about claims or rumours that DNA from Russians could be used to build some kind of 'genetic weapon'. In June, the government ordered that Russian DNA must no longer leave the country. But the order was revoked just two weeks later after an outcry by scientists and the media. Severinov adds that a newly emerging anti-Americanism, and widespread aversion to Western scientific culture and journal editors are also major obstacles on the road to reform. More money will not help without proper peer review and competition, he says. "Too many people here think that grant writing is just a detraction from daydreaming and doing 'wonderful' science."

Scientists, such as Severinov, who return to Russia, are still rare. A special federal programme for Human Resources in Science and Technology, aimed at attracting and enticing young people back to scientific careers, is to be launched in 2009. But Sobolev fears that Russian researchers will continue to go abroad, and stay there, because they have no chance of starting an independent and reasonably well-paid group in their home country. "I am not very optimistic that things will improve soon," he says. "The scientific establishment is still opposed to true reform, and the next generation is absent. We can only hope that many young scientists will return one day and proceed to the political level." In the long term, he says, there is no way to escape the necessary changes: "Isolation is not an option." ■

Quirin Schiermeier is a reporter for *Nature* based in Munich.

See Editorial, page 507; News Feature, page 528; and Commentary, page 536.

©2007 Nature Publishing Group

What the scientists say

Russian researchers, and those who have worked in Russia, share their thoughts with *Nature* on the problems faced by the country's scientific system — and how they could be addressed.

Alexander Sobolev
Geochemist, Vernadsky Institute of Geochemistry and Analytical Chemistry, Russian Academy of Sciences (RAS), Moscow. Corresponding member, RAS

According to the citation-analysis company Thomson Scientific, Russia is eighteenth among countries ranked by citations in the scientific literature over the past 10 years. That is a result not just of low overall funding but because management of basic science still stands on the concepts of a closed society, with a centralized administration inherited from the days of the Soviet Union. This leads to the absence of international peer review and to little motivation for scientists to produce international-level scientific results — they do not really need them to get funding from national sources. In addition, centralized funding of institutions, rather than of individual scientists, leads to resources being wasted.

"Russia needs a nationwide initiative to restore its scientific potential."

Compared with the Soviet Union, Russia today is quite an open society and has had to adjust its national science to the international level. But so far only timid steps have been made in this direction. The Russian Foundation for Basic Research, set up in 1992, is still almost the only source of peer-reviewed funding for individual natural scientists in the country. It suffers from a shortage of competent experts, and provides insufficient support — up to only US$15,000 per year per project.

I think Russia needs a nationwide initiative, concentrating on personnel rather than specific technological goals, to restore its scientific potential. I would start by establishing an international peer-review expert system. The core of this should be formed by Russian scientists — both inside and outside the country — who are producing internationally recognized results.

The next step would be to concentrate funding on individual scientists and groups who operate at the international level. This should include open competition for adequately funded permanent positions and 'soft money' grants based on international peer review. It would also mean creating 'centres of excellence' to support efficient groups with infrastructure and equipment. These ideas are already being discussed by Russia's scientific community (see www.scientific.ru), and I hope they will finally be considered by the authorities.

Konstantin Severinov
Waksman Institute of Microbiology, Rutgers University, New Jersey. Institute of Molecular Genetics, Russian Academy of Sciences, Moscow

Russian biological science has a proud history, with greats such as Ivan Pavlov and Nikolai Koltsov, and successful emigrés including George Gamow and Theodore Dobzhansky. Despite Stalin's purges of the 1930s and 1940s, by the 1970s the country had the nucleus of a thriving life-sciences community. A precipitous decline in funding conditions in the 1990s caused many active scientists to leave the country. This brain drain is only now beginning to slow, thanks to the recent flush of petrodollars, but significant problems remain.

Basic research in biology is deemed unnecessary by Russian society, even more so now that scientists' salaries are starting to increase and society expects immediate practical results in return. Russia has no modern biotechnology industry to speak of, so there is little demand to exploit basic-science results. A scientific superstar might raise the profile of basic research. But without a vibrant scientific community, and with problems in getting reagents and sending materials abroad, such a star is unlikely to emerge in Russia.

Nevertheless, the nostalgic view of great Russian science is deeply ingrained in the public mind. To return to this greatness, the government has developed grandiose programmes such as the multibillion-dollar nanotechnology initiative. Although welcome, this influx of money may also make science attractive to profiteers doing bad (or pseudo) science.

Certain survivors of the 1990s who stayed in the country and remained in science are now fashioning themselves as 'keepers of the flame'. They are moving to top administrative positions in the institutes. Some tend to be conservative and against the use of bibliometric data to measure productivity. Some propagate the myth that editors of international journals are prejudiced against Russian scientists. As a result, a significant number of students who stay in the country end up doing science of dubious value.

At least one programme at the Russian Academy of Sciences, in molecular and cellular biology, is giving generous three- and five-year grants to individual groups, including new groups set up by returning Russian scientists. But establishing financially and scientifically independent labs in academy institutes goes against the general idea of centralization in present-day Russia — so it is unclear how long such labs will last.

Q.SCHIERMEIER

Mikhail Kovalchuk
Director, Russian Research Centre Kurchatov Institute, Moscow. Acting vice-president the Russian Academy of Sciences (RAS). Corresponding member, RAS

Over the decades, Russia has initiated a lot of important fields in world science and technology, including atomic energy, nuclear fusion and space exploration. Despite the financial troubles of the past 20 years, science in Russia has continued to develop, thanks to the intellectual and material reserves from the Soviet period. Our research has become integrated in the world scientific community.

"Russian research has become integrated in the world scientific community."

Examination of current world science has highlighted new targets for scientific and technical development, which has led Russia to rethink its approach and has generated a new state policy on science and technology.

World science now is characterized by a movement towards technologies of atomic or

©2007 Nature Publishing Group

molecular design, combining modern microelectronics with biological structures. This merging of disciplines needs to be reflected in the scientific expertise. Russia's present science and education system, which fosters narrow specialization, must be replaced by a system based on interdisciplinary research.

The simplest and most effective model for such a development would be to set up national scientific laboratories. In Russia this is now being piloted with the nanotechnology project at the Kurchatov Institute. An international network of such research centres, based on national laboratories, could be the foundation of a new global system of scientific and educational organization.

Mikhail Feigel'man

Physicist, L. D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow

There are five big problems for fundamental research in Russia. First, the salaries of researchers are still far too low: the official salary of a full professor at an academic research institute in Moscow is about one-third of the average income in the city. Second, although funds for modern research equipment have, in principle, become available over the past few years, distribution of this money is irresponsible — it usually goes to 'big shots' not doing real science. Third, the top brass of the Russian Academy of Sciences has been unable to start long-anticipated reforms of its own organization. Fourth, government officials responsible for reconstructing the scientific infrastructure have often been simply incompetent. And fifth, the federal secret service has accused numerous Russian scientists of espionage, with obvious falsifications.

The result is that the number of internationally competitive groups has seriously diminished in the past 15 years. The few that have survived could potentially seed a revival of physics in Russia. The key problem is the lack of a generally recognized system of research evaluation. There is no well defined procedure to distinguish really good science from pure trash.

What can be done for recovery? First, the system of research evaluation can and must be constructed by the scientific community itself, using the very modest financial means provided by non-governmental sources. Second, Russian business interests could be approached for help, as they want to develop a national high-tech industry. Third, real face-to-face international cooperation between specific labs and between scientists in Russia and those abroad should be encouraged as much as possible.

Taking all that together, we can probably survive some more years — with hope for a more adequate Russian government to come, eventually, and to do its job in rebuilding the national science system.

Rebecca Frumkina

Institute of Linguistics, Russian Academy of Sciences, Moscow

Although I am more than 75 years old, I am still an active academic. This means that I am considered socially successful and, by all formal standards, I should have nothing to complain about. In fact, I am deeply concerned about the future of Russian science — and here I use the term 'science' to encompass both science and the humanities.

Science is a social institution. As such, it has its own laws and specific means of interaction with other social institutions. So science should not be treated as an enterprise that needs an external crisis manager to make ends meet. Nevertheless, this is exactly what state policy does — examples include attempts to quash even the slightest signs of autonomy at the Russian Academy of Sciences; the reduction of money for fundamental research; and the cutbacks and neglect at the National Library.

Young people who wish to become academics will go on leaving this country until the state changes its position on fundamental aspects of life such as conscription into military service and housing. It is well known that Russia's academics are underpaid, so I am not surprised that nearly all my former postgraduates either left science here for, say, the media or advertising, or have left this country to join companies such as Microsoft, Hewlett-Packard or IBM.

"I am deeply concerned about the future of Russian science."

I reject the prospect of obligatory EGE exams for students attempting to enter university, as well as Russia's plans to join the Bologna process, which aims to standardize university degrees across Europe. In Russia we do not have a stable education system leading to formal degrees such as a BA, MA, BSc or MSc. We need to have a stable system in place before we can be ready to standardize our academic degrees. Otherwise, all efforts will be counterproductive.

K. G. TOOHEY

Jens Kuhn

German virologist, New England Primate Research Center, Massachusetts

Russian virology is not typically held in high esteem by most Western scientists. Almost certainly this is because in the West, the quality of research is mainly defined by the impact factor ascribed to the publishing journal. Consequently, articles published in journals with low impact factors, such as the majority of Russian virology contributions, are considered mediocre science in the United States and Europe. This assessment is unjustly biased.

Russian virological data are frequently published in journals not indexed in PubMed, which, sadly, is often the only database used by Western scientists. Political, rather than scientific, factors — perhaps patriotism or the demands of the still ever-present state-security apparatus — often drive Russians to publish their data in national journals in the Russian language. Furthermore, publication in a journal with a high impact factor is often facilitated by state-of-the-art technology, which can rarely be afforded by Russian laboratories. Limited finances may mean a shortage of reagents and outdated equipment. But taking a careful look at what is being published under these dire conditions, one can only be impressed by the skills and creativity of many Russian researchers. For instance, my search of libraries has shown that the majority of the world's scientific literature on both Crimean-Congo and Omsk haemorrhagic fevers has been published in Russian, but most of it has not been indexed in PubMed.

A. YARKHO

It is true that much of this literature is not cutting-edge molecular biology — but in terms of epidemiological, clinical and pathological data, Russian science equals literature from Western labs. As for molecular-biological data, there are excellent and pioneering publications devoted to genomic sequencing of tick-borne encephalitis, Marburg and smallpox viruses. Of course, it should not be forgotten that, especially in pathogen research, good-quality Russian research often sprouted from 'bad' circumstances — namely, the Soviet Union's biowarfare programme, which was terminated more than a decade ago. Various international grant-assistance programmes are helping to redirect former biowarfare facilities into public health; and international scientist exchanges, such as the one in which I participated in Russia, are boosting this process (see *Nature* **423,** 679; 2003).

There is much to be learnt from Russian researchers, and the West should do all it can

©2007 Nature Publishing Group

to support peaceful Russian science both financially and collaboratively — rather than be arrogant and ignore its important contributions.

Roald Sagdeev
Physicist, University of Maryland.
Full member, Russian Academy of Sciences

As we commemorate the fiftieth anniversary of Sputnik, should we feel nostalgic for the past grandeur of Soviet science? Sputnik's historic launch took place only a year after Stalin's cult of personality was denounced by Khrushchev. Science back then was simply a servant of the regime. Research areas related to defence received a disproportionate amount of state support; the rest, at best, was on a sort of welfare. The government invested in expanding the infrastructure by building science cities (akademgorodoks). Khrushchev's successor, Brezhnev, declared science to be a "productive force of society". But the productivity of researchers was limited by factual isolation from international science, and by tight ideological and administrative control.

The dismantling of the Soviet Union has brought tangible fruits of globalization to the Russian scientific community. Researchers now have the chance to become part of world science and to look for work outside the country. But, ironically, the dividends of freedom have not led to many more opportunities at home.

If the impediment to science investment in the early 1990s was coping with the economic cataclysm of the post-communist transition, today it is the need to adapt to the difficult realities of the market economy. That requires deep institutional and structural changes in the organization of Russia's research and development; changes that should see funding distributed fairly to researchers across the country — the real 'productive forces' — rather than being made available to only a privileged few. Fundamental research and large projects in applied science will need stronger support from the state, similar to that already provided for Russia's participation in the international fusion-energy experiment, ITER, and now promised for nanotechnology.

Contemporary science culture is quite cosmopolitan. Scientists are choosing the places most conducive to the application of their talents. Reforms that provide real, rather than simply declared on paper, successes in Russia are needed to bring back her estranged scientists. By doing so, Russia could perhaps achieve even more and attract the best foreign brains. Wasn't such a thing done in the eighteenth century by Peter the Great? After all, today's Russia is richer and already allows herself to recruit expensive foreign sports stars.

Simon Lyakhovich
Theoretical physicist, Tomsk State University

The period after the collapse of the Soviet Union can be divided into two stages. From 1992 to 2002, a new system was established that was characterized by calls for research projects or grants that would be awarded on the basis of peer review. During this turbulent period, resources were limited, and the social environment was not conducive to doing fundamental research. As a result, individuals would typically make extended research visits to the West. But what is rarely recognized is that the opinion of the scientific community mattered in this system, and that scientific excellence was a criterion in selecting which projects or teams received support. Even with limited resources, shaky institutions, immature mechanisms, and multiple mistakes and abuses, this system delivered some impressive results.

Starting around 2003, the research-oriented team system was replaced by an administration-oriented one, based on new oil and gas wealth, and built on the new social order in the country. Organizations and administrative ranks are now the key aspects. Russian funding sources are growing rapidly, but the funds are now directed to organizations far more than to research teams. Selection decisions on the huge grants are made through negotiations with the director of the host organization rather than through serious independent evaluation of the scientific quality of the proposal and the research team. And location now plays an increasing role: in the extremely centralized administrative system, Moscow and St Petersburg scientists have better opportunities, on average, than their colleagues elsewhere.

The flow of talent to the best universities outside these two cities has begun to decrease. But the cost of attracting student talent is not high. For example, the 20 best universities could be boosted by establishing 100 stipends per university to attract talented students from other cities. These stipends would each cost only US$400 per month — so the cost of the overall programme would be almost nothing compared with the funds spent on the inefficient mega projects of the new era.

Evgeny Antipov
Chemist, Moscow State University

The biggest challenge facing the Russian scientific community is the government's attempt to start an immediate breakthrough in innovative technologies by significantly increasing funding. This move may sound very promising, but after 20 years of financial starvation our scientific community has become very inhomogeneous. Only a few groups have been able to maintain the universally accepted international research level and can win competitive international grants.

The existing funding system in Russia is inefficient and ignores international standards. The government's desire to move beyond development based on natural resources will not be realized with the simple formula 'more money = better science'. What needs to be done is to establish trusted scientific expertise to evaluate the quality of research and therefore the fields of science and groups to be funded. That expertise would be based on accepted international criteria, such as quality of publications, citations and scientific reputation in the international community. It would modify the formula to be 'more money = better scientific expertise = better science'. Such an approach would stop the support of low-quality research, which creates no real base for innovation.

V. EGIKOVA

I see another challenge in establishing solid links with fast-developing private companies. These have already started investing to develop new products and to improve existing technologies, but have encountered the same problem: how to identify people who can do high-quality research. Many of them prefer to buy advanced technologies from abroad; fortunately, some companies continue their search for scientists in Russia. For instance, Moscow State University and the United Company RUSAL (one of the main aluminium producers) last year established a joint lab to develop the basics of aluminium production. It was one step towards the 'revival' of our science — I hope it is not the last. ■

See also pages 507 and 536.

©2007 Nature Publishing Group

B. LUNDBERG/NATUREPL.COM

A QUESTION OF SURVIVAL

International collaboration and a can-do spirit have allowed some Russian scientists to flourish. **Alison Abbott** watches an extraordinary field test for mutant mice in the Russian wilderness.

As she watched Mikhail Gorbachev's resignation speech on Christmas day 1991, Inga Poletaeva, a Russian behavioural geneticist, was in Switzerland, on her first visit to the West. For her, as for most Russians, the fall of the Soviet Union seemed inevitable. Unfortunately, the same fate also seemed imminent for her research. Poletaeva was a guest of Hans-Peter Lipp, a Zurich University neuroscientist to whom she had sent a forlorn letter a year earlier.

Poletaeva was looking to forge a collaboration, applying Lipp's neurobiological methods to strains of mice bred at Moscow State University. Even with perestroika in full swing, science funding had collapsed and prospects were grim. Poletaeva's department didn't have enough money to feed its mice, let alone experiment on them.

During that short visit, she and Lipp became firm friends. Talking science into the night, they laid plans for a project — correlating the size and distribution of one sort of brain cell, the mossy fibres of the hippocampus, with the behaviour of her mice. As conversation drifted on, Poletaeva mentioned her friend Valentin Pazhetnov, who had set up an outpost at a place called Bubonizi, deep in the Russian Tver forest, for raising and releasing abandoned baby bears. Her words, she recalls, "were like a trigger".

Lipp had a problem with the direction behavioural genetics was taking. The 1980s revolution in genetic engineering had made it easy to produce mutant mice — by targeting a single gene and knocking it out of the genome, for example, or transferring a new one in. New mice were being created, and they would emerge from the standard battery of lab-based behavioural tests with labels such as 'smarter', 'anxious', 'depressed' or 'psychotic'. But did these laboratory measurements have any real-life meaning, he wondered? Bubonizi sounded like just the sort of place to answer this question.

Lipp, still restless today at 60, says he saw the opportunity to test mice in wild conditions in Russia, something that would never fly in his home. "I would never embark on anything that was against Swiss laws or European Union rules — or Russian rules," says Lipp. But he recognized how time-consuming and exhausting it would be to try and set up such a station in Switzerland, where popular movements against genetic engineering regularly delayed research plans.

The outpost

Poletaeva took Lipp to visit Pazhetnov that summer, the only time of year such trips are easy. Bubonizi is about 500 kilometres northwest of Moscow — an 8-hour drive on the narrow, pot-holed road to Riga, where the only signs of civilization are the occasional road-side stalls where peasants sell berries, honey and wood-smoked river fish.

Arriving at the outpost, and with Poletaeva acting as interpreter, Lipp lost no time in arranging for Pazhetnov and his family to organize the building of the first animal house there. He put US$500 of his own cash on the table to cover the costs. It was a risk. "I couldn't know whether my money would evaporate into vodka or anything else," he says.

When he returned the next year, Lipp was pleased to find a splendid building, constructed from local logs. Over the next couple of years, other facilities were added. Pazhetnov's people built a log cabin field laboratory where mouse brains can be analysed histologically to

©2007 Nature Publishing Group

determine neurobiological differences between mice coping well or badly with the great outdoors. Upstairs they installed simple living quarters. Old furniture, discarded lab equipment and computers brought in from the West were supplemented with whatever could be found locally. An old vodka still, for example, was resurrected to distil water.

The facility's centrepiece is eight pens the size of basketball courts that spread out over a large clearing. Looking somewhat akin to overgrown swimming pools, they are fortresses designed to mimic life on the wild side, while still keeping things monitored, controlled and secure. Their metre-tall concrete walls are dug deep into the ground to prevent the mice from tunnelling out and forest animals from tunnelling in.

Electric fences atop the walls stave off larger animals and in the unlikely event that mice escape, they would have a long way to travel to reach the kind of human settlements in which they could thrive or multiply. The pens are equipped with small wooden shelters where the mice nest. The environment closely mimics the real world of wild mice except for one thing — they don't have to forage for food. It is provided for them at covered plates set down at strategic distances from the shelter and fitted with sensors to detect the individually coded transponders inserted under the skin of the mice. Narrow wooden covers run between the shelters and the plates. Without them, owls would easily pick off exposed subjects on their way to dinner.

By changing the timing and placement of food, scientists can examine, for example, how nervous or curious mice are, or how good they are at remembering, or predicting, where the food is going to be. When a transponder no longer shows up at any of the plates, it's assumed that the mouse has either died or become a meal for an owl.

Over the years, Lipp has run more than a dozen experiments with Poletaeva — whose research group, she says, was saved by Lipp — and other collaborators around the globe. Short experiments can be run to follow behaviours in the pen, and longer experiments, sometimes lasting years, can look at the changing genetics in a mixed population of mutant and wild-type mice.

"It may be that the most important thing for life in the wild is not so much smart brains, as respect for danger." — Hans-Peter Lipp

Different in the real world?

One of Lipp's first long-term experiments was on mice that were not mutant, but bred. He began by looking at mossy fibres in the hippocampus, a brain area associated with learning and memory. Projections from these fibres reach up from one part of the hippocampus to form synapses with neurons in another part. Many studies in different labs had shown that the length of these projections correlate positively with learning performance in very specific lab-based tests, such as different types of mazes. "I wondered if natural selection would favour the longer mossy fibres," says Lipp, "or if those very specific lab tests were misleading in terms of what really matters in the 'real world.'"

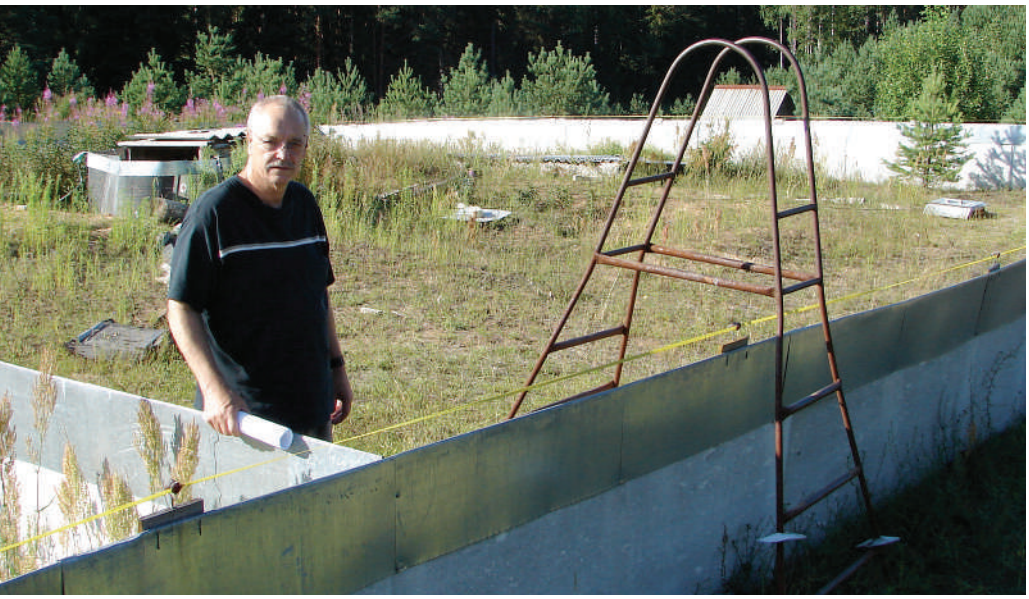
So Lipp and Poletaeva bred four strains by crossing mice that had long hippocampal mossy fibres with those that had short fibres. They divided the offspring into three independent 'founder' populations in which each mouse inherited one gene for long fibres and one for short fibres. One founder population remained in the lab, breeding randomly. The other two made their way down the bumpy road to Bubonizi where Lipp placed them in separate outdoor pens. There they bred for more than four years.

With each generation change, around three per year, Lipp and his postdocs caught the mice, selected a random sample for brain histology and released the others. Members of Pazhetnov's family helped look after the experiments when the scientists dispersed for the harsh winters.

During the fifth winter, the population in both pens mysteriously crashed, a phenomenon that Lipp says has played out again and again in long-term experiments. Nevertheless, with 12 generations' worth of data still being analysed, the emerging trends were disconcerting. Ongoing analysis seemed to suggest that natural selection in the 'wild' was weeding out the long-mossy-fibred animals, not the short-fibred ones. The control founder population breeding in the lab in Moscow had not changed.

The trend was visible within a year. When Lipp put mice from each pen through the normal routine of laboratory tests, they seemed to be more timid and less inclined to explore than the control, lab-bred mice. Nevertheless, they performed about as well in learning and memory tests. When the experiment was repeated in fully covered outdoor pens, safe from predators, mossy-fibre length did not change through the generations. Lipp says he thinks that earlier studies on learning and mossy-fibre length may have underestimated the impact of this trait on exploratory behaviour. "It may be that the most important thing for life in the wild is not so much smart brains, as respect for danger," says Lipp. "Maybe the mice with longer fibres were being picked off by owls because they were less afraid to leave familiar, sheltered areas." If the complete analysis of his data confirms the trends, he says, it implies a shift to defensive behaviours.

Mutant 'smart' mice may make interesting test cases. In 2002 behavioural geneticist Shigeyoshi Itohara's group from RIKEN Brain Science Institute in Saitama, Japan, engineered mice to lack the *S100b* gene. Previous studies had shown that overexpression of the gene limited a neuronal process called long-term potentiation, which has been correlated to learning and memory. Itohara's knockout mice have enhanced long-term potentiation in the hippocampus



A. ABBOTT

Hans-Peter Lipp in a pen that allows him to monitor mice in a close-to-wild environment.

©2007 Nature Publishing Group

and perform well in spatial memory and fear conditioning tests (H. Nishiyama, T. Knopfel, S. Endo and S. Itohara. *Proc. Natl Acad. Sci. USA* **99,** 4037–4042; 2002).

Similarly a group led by Hee-Sup Shin, a behavioural geneticist from the Korea Institute of Science and Technology in Seoul, has bred mice that lack a neuronal membrane protein called NCX that controls levels of calcium in neurons and thus their electrical activity (D. Jeon *et al. Neuron* **38,** 965–976; 2003). Both types of mice perform better in laboratory maze tests, learning their way around faster than their wildtype counterparts. They are potentially very interesting tools for the learning and memory community.

H.-P. LIPP

A. ABBOTT

In the woods: Russian scientists work inside the isolated field station (left).

More brains than sense

Lipp says he thinks that in the wild most smart mice will be more active, and that activity will be their undoing. Indeed, a short-term study over a few weeks this summer, which challenged the intelligence of the Japanese mutant by changing positions of eight food plates, seemed to suggest that they were more active than the wildtype in seeking out the plates, but not very precise at remembering where they all were.

Long-term studies on the Korean group's mice will commence next year, and whereas Lipp sees a foregone conclusion, Shin says the potential is clear. "My NCX mice learn quicker how to find their way to the submerged platform in a water maze and other tests of spatial learning and memory," he says, "so they are smart in the way that London taxi drivers are — but would that smartness help them survive in the real world?" Only experiments at Bubonizi can help, he says.

In addition to his mice currently in the field, Itohara wants to test mice that seem in lab tests to be unaffected by their mutations. "Maybe there are subtle phenotypes which we don't see in our battery of tests looking for very specific deficits — which is after all a very limited way to understand mouse brain function."

But Bubonizi has its own limitations. "The only questions we try to answer here are to do with biological fitness and learning — does the genetic change have any relevance for real life?" says Lipp. Although the set-up is in some ways contrived, even coming close to 'real life' means contending with nature's unpredictable whims. Aside from the mysterious crashing of mouse populations in the pens after four or five years, uninvited guests have beaten the tight security. One particularly brutal winter, a snow drift breached a pen wall allowing a mink to climb in and eat all the mice.

Logistics are a problem as well. It's a long and rough journey from Moscow — and it's also a bureaucratic nightmare to bring mice into the country from abroad. Getting all the right forms has delayed some experiments for up to a year. That's not to say things never work: Lipp was amused to receive a consignment of Itohara's mice from a world courier service that drove right up to the Bubonizi log huts. "Really just like the television advertisements where the courier shows up at the most unlikely place at the end of the Earth saying 'we deliver anywhere!'"

Time running out

This is a far cry from the early days of the station when Poletaeva insisted Lipp was always accompanied on the drive from Zurich for fear of the Russian mafia. He would load a two-tonne truck with all the necessary items, lab equipment, chemicals, Italian wine and of course Swiss chocolate. He got used to the long drives but never to border checks, which he says could last from hours to days, often requiring a bit of cash to ease passage.

The trip remains impossibly tiring. Serge Daan, a behavioural biologist from the University of Groningen in the Netherlands, is looking at circadian rhythms using a popular mutant that is missing a fundamental clock gene, *Per2*. He has visited Bubonizi three times, and enthuses that the people are "so helpful and reliable". Regardless, he won't be returning. Together with postdoc Kamiel Spoelstra he is setting up a similar station a 15-minute drive from Princeton University in New Jersey where he has teamed up with Martin Wikelski, an expert in telemetry from Princeton. Using this technology, he will be able to monitor the movement of his mice fitted with telemetry chips 24 hours a day. "Our station is very much inspired by Bubonizi," he says. "But the technology system is better, and the infrastructure for scientific back-up is more secure." And most of all, it is easier to get to. "There is a limit to how many times you can make that forbidding trip."

"I wonder if there will be a biological price to pay in the real world."
— Hans-Peter Lipp

According to Swiss regulations, Lipp will have to retire from his university in five years' time and the fate of the station thereafter is uncertain. Its log housing won't last without extensive repair work for which he is unlikely to be able to raise money. "Maybe someone will take it over — or maybe I'll use it as a holiday dacha," he says, using the Russian word for country home. Otherwise it will be abandoned to the forest and Pazhetnov's bears. But when contemplating retirement and Bubonizi's end of life, Lipp is already thinking about the experiments still to do. One burning wish: testing mutant mice that have shown remarkable longevity. "I wonder if there will be a biological price to pay in the real world," he says. ■

Alison Abbott is *Nature*'s senior European correspondent.

©2007 Nature Publishing Group

Biometrics: still much too unreliable for everyday use

SIR — Anil K. Jain gives a misleading picture of the effectiveness of biometric technologies, which identify people on the basis of physical features (News & Views Q&A 'Biometric recognition' *Nature* **449,** 38–40; 2007).

Jain says that, since the inception of the US Visitor and Immigrant Status Indicator Technology (US-VISIT) in January 2004, more than 75 million visitors have been processed through it and about 1,000 have been denied entry. According to the US Department of Justice's analysis of US-VISIT after its first year of operation (www.usdoj.gov/oig/reports/plus/e0501/exec.htm), on average 118,000 people pass through each day. Of these, 22,350 are subject to secondary inspection and 1,811 of these are refused entry — considerably more than Jain states.

Of those who are subject to secondary inspection, 92% are subsequently admitted to the United States. These figures raise doubts about the reliability of the biometric systems used for primary inspections.

Applications such as US-VISIT need to reject virtually all impostors, and keeping the false accept rate (FAR) close to zero pushes the false reject rate (FRR) up. Jain states that current fingerprint-recognition systems can provide a FRR of up to 0.01% at a FAR of 0.1%. However, in a 2005 UK Passport Service trial (see http://dematerialisedid.com/PDFs/UKPSBiometrics_Enrolment_Trial_Report.pdf), the FRR for fingerprints was 19%.

One objective of the proposed UK National Identity Scheme is to make it easier for people to prove their identity. At current performance, biometrics based on fingerprints could instead make it harder for nearly a fifth of the population to prove that they have the right to work in the United Kingdom and enjoy social entitlements.

David Moss
Business Consultancy Services Ltd,
58 Vineyard Hill Road, London SW19 7JH, UK
http://dematerialisedid.com

Biometrics: easy to steal, hard to regain identity

SIR — In your News & Views Q&A 'Biometric recognition' (*Nature* **449,** 38–40; 2007), Anil K. Jain asserts that biometric technologies are more difficult to abuse than traditional methods of identification. However, we all leave a biometric trail in our daily lives: our fingerprints on a drinking glass, our voice on a telephone answering machine, our iris patterns on a photograph. We have little ability to change such characteristics and little control over this trail, which makes biometrics useful to forensic science.

But it's exactly these properties that make biometrics a poor replacement for passwords and ID cards, since it's easy for an intruder to collect someone's fingerprint or iris scan without their knowledge, and then inject it into a biometric identification system. Even if the victim becomes aware of the problem, it's impossible to revoke the biometric. If your credit card is stolen, the card company can send you a new one with a different number, but you can't get a new set of fingerprints.

It is precisely because biometric information is irrevocable and unwittingly provided in our daily lives that it is so useful to organizations that regulate the individual (for example, the US Immigration Service), but of little use where the individual controls identification and authorization.

Andrew Watson
72 Kimberley Road, Cambridge CB4 1HJ, UK

Organic synthesis remains relevant in drug discovery

SIR — As you point out in your News story 'Chemists synthesize a natural-born killer' (*Nature* **448,** 630–631; 2007), some researchers question the merits of organic synthesis, whereby chemists seek to recreate a synthetic version of a natural product. In its defence, I want to remind readers that when Elias James Corey received a Nobel prize in 1990 for the development of organic synthesis, the press release stated: "To perform the total syntheses successfully, Corey was also obliged to develop some fifty entirely new or considerably improved synthesis reactions or reagents…which…have become commonplace in the synthesizing laboratory."

Such newly discovered reactions and reagents are routinely used in the discovery of drug candidates. The cost of pharmaceuticals is, to some extent, directly related to the ease or complexity of their synthesis, and there are still many drugs in the pipeline whose development has been hindered by problems encountered during their manufacture. This underlines the point that the science of organic synthesis is still a worthy cause to be pursued.

Mukund M. Mehrotra
Portola Pharmaceuticals, 270 East Grand Avenue, South San Francisco, California 94080, USA

Ocean-drilling vessel should soon be afloat again

SIR — Those involved in the Integrated Ocean Drilling Program (IODP) are delighted by your encouragement of their work in your Editorial 'Drill often, drill deep' (*Nature* **449,** 260; 2007) and News Feature 'Staying afloat' (*Nature* **449,** 280; 2007). However, it is not correct to state that the IODP's US platform, the *JOIDES Resolution*, has languished at a shipyard since 2003. This research vessel completed 10 expeditions between June 2004 and December 2005, each with a full complement of scientists from the United States, Europe and Japan.

The *JOIDES Resolution* has been in a Singapore shipyard since September 2006, undergoing a complete overhaul. Work has continued since it entered the shipyard, although the hostile business climate described in your Editorial has caused a few delays, as sky-high oil prices have meant higher costs and increased pressures for time in the yard. Interested readers can track progress via pictures of the overhaul at www.joiscience.org/SODV/status.

The *JOIDES Resolution* is scheduled to resume drilling activities in spring 2008. The shipyard schedule, weather conditions and Japanese fishing-fleet restrictions will not allow it to join its sister ship, the Japanese *Chikyu*, in 2008. But the *JOIDES Resolution* will be an active participant in NanTroSEIZE, the Nankai Trough Seismogenic Zone Experiment, by 2010.

Steve Bohlen
Joint Oceanographic Institutions,
1201 New York Avenue NW,
Washington DC 20005, USA

Sic is not enough; historical wording should be changed

SIR — Your decision to insert *sic* after "scientific men" in reprints of your historical mission statement ('Men [*sic*]' *Nature* **448,** 728; 2007) shows that you are more concerned about the historical integrity of your mission statement than for decency and justice. Why make a mere "tiny step" in the "right direction"? A simple addition of "…and women" would have brought you to the goal.

Isn't it enough that even such a "tiny step" took *Nature* nearly 140 years? Pious traditionalism is being used to excuse prejudices in many quarters, but it is shocking to encounter it in a supposed bastion of enlightenment and intellectual progress. Do not consider yourself excused.

Sanya Samac
Imaging Research, Sunnybrook Research Institute,
2075 Bayview Avenue, North York,
Ontario M4N 3M5, Canada

The Editorial related only to reprinting of our historical 1869 mission statement, not to our current statement (www.nature.com/nature/about/index.html), which uses the term 'scientists'. Comments are welcome at Nautilus, http://blogs.nature.com/nautilus/2007/08/women_and_men_of_science.html — Editor, *Nature*.

©2007 Nature Publishing Group

COMMENTARY

In 1988 the unmanned spacecraft Buran (left) orbited Earth and landed successfully before the project was cancelled. It met its demise in a hangar collapse in 2002.

R. RESSMEYER/CORBIS

Breaking up is hard to do

Economist and former science minister **Boris G. Saltykov** sees opportunity for change and innovation in Russia.

Since the break-up of the Soviet Union, the Russian government has often declared a desire to create a 'knowledge-based economy'. In 2007, there are signs that a type of national innovation system can be created in Russia, assuming that a series of important institutional reforms are adopted. How will these moves affect the balance of power in Russian research and development (R&D)? To answer this question we need to understand the condition of Russian science at the beginning of the twenty-first century.

The fiftieth anniversary of the launch of the first Sputnik satellite reminds us of a time when two great powers — the United States and the Soviet Union — began an exhausting technological race. This was the symbolic beginning of a 'golden age' of Soviet science (see page 542) during the 1960s and 1970s, with a long line of Russian achievements in space, aviation and nuclear physics, the creation of unique research instruments, including the world's first tokamak — a fusion reactor — and world-class facilities, such as the Joint Institute for Nuclear Research in Dubna.

But Soviet science and technology (S&T) was based on a paternalistic model. All resources were centrally distributed and the highest priority was given to military-related research. There was weak communication between sectors, with each industrial research centre reporting to its branch ministry, and with most basic research under the control of the independent Russian Academy of Sciences (RAS) in Moscow. Little research was conducted by universities.

The Soviet system also created dozens of 'science cities', built in remote areas or in secure compounds alongside civilian cities. In all cases, there was little interaction between scientific–technical organizations and the wider Soviet economy. And of course all assets, including intellectual property, belonged to the state.

Such a model allowed the state to concentrate huge resources on a restricted number of prioritized S&T projects, such as Sputnik, and allowed basic research to flourish. In 1990 Soviet science was still a heavyweight enterprise supported by high levels of public funding, almost 2 million highly trained scientists, engineers and technicians and more than 4,600 highly specialized research institutes. But the system was wasteful, conservative and unable to respond quickly to new challenges. And in the long term, there was no effective way of renewing the ranks of élite researchers. The isolated and specialized institutes had little chance of adapting successfully to life after the end of Soviet rule.

The last major success of Soviet science was the launch in 1988 of the reusable unmanned spacecraft Buran, which orbited Earth twice before landing at Baikonaur Cosmodrome, Kazakhstan. The project — begun in the 1970s to rival the US space shuttle programme — was the most expensive in the history of Soviet space exploration. It was cancelled in 1993, with one flight to its name, because the Soviet economy could not continue to supply the necessary resources.

Economic crisis

After the dissolution of the Soviet Union in 1991, the new government of Russia needed to solve two problems: to prevent the complete collapse of the economy and to move towards a market-based system. Moves to attain this had dramatic effects on Russian science: by 1992 the R&D budget was three to four times smaller than it had been in 1990, industrial demand for R&D had vanished, and by 1995 the number of civilian R&D personnel — half of whom were researchers — dropped to around a million. Much of this contraction has been attributed to poor wages causing younger researchers to migrate to jobs abroad or switch to more lucrative professions within Russia.

Three steps were needed for R&D reform: demilitarization, integration into the world research community and democratization of R&D management. The huge shortage of resources dictated two main strategies: to support only the best teams and to give money to priority fields of research. This approach required new financing mechanisms and structural changes across the R&D system.

During the first stage of reform (1992–96), when I was science minister, four new 'targeted

©2007 Nature Publishing Group

foundations' were established, including the Foundation for Basic Research. These agencies first introduced the concept of competitive funding to Russian research. At the same time, civilian industrial-research institutes and design bureaus were privatized. New legislation for science, including modern patent law, was adopted. But these sweeping reforms did not touch the other two sectors of public research — the RAS and the military S&T sector.

Since the Soviet collapse the RAS has focused on preserving the status quo (see page 524). Because of this strategy, and the contraction in military and industrial spending, the share of basic sciences within the federal R&D budget rose from 10% in 1992 to 30–35% in 1996, although it declined by three to four times in real terms. The number of researchers also dropped while the number of RAS institutes grew, from 330 to 450 by 2004.

In general, the number of R&D organizations declined across Russia, falling to 3,500 by 2005. Others switched to commercial activities unrelated to science — leasing land or property to business — to survive the economic crises of the 1990s. The population of Russian scientists is also ageing quickly: by 2004 researchers in the most productive age group, 30–45 years, had almost disappeared and the average age of professors exceeded 60.

Publishing productivity noticeably diminished and Russia dropped from fourth or fifth place in the world hierarchy in 1990 to eleventh place by 2005. Publications by Russian researchers in leading scientific journals declined by 20% over the 10 years to 2005, with the share of Russian articles in world publications falling from 3.64% to 2.26% (ref. 1).

M. LYSTSEVA/ITAR-TASS

Future hope: a scale model of Superjet 100, a new Russian jet built with international partners.

A new lease of life

When Russia's economy finally recovered after the shocks of the 1990s, it was heavily dependent on exporting natural resources — oil, gas and metals. But after the election of Vladimir Putin as the president of Russia in 2000, the government turned to S&T as a means of diversifying the economy.

"Only a new, young élite can help Russia become a modern knowledge economy."

Following the re-election of Putin in 2004, the modernizing activities of the Ministry of Education and Science increased sharply. Today's ministry has two important advantages compared with that of the early 1990s: a larger R&D budget and a strong team close to the president. The new minister, Andrei Fursenko, has prepared a strategy for the development of Russian science until 2015 and initiated a federal targeted programme in priority areas of R&D. Similar to the targeted foundations, this programme allows competitive project funding, and so is helping to shift Russian science away from basic institutional funding.

The old élite of the RAS have resisted the ministry's attempts to reform it, so, in 2007, the ministry made a strong and unusual move — they established the Russian Nanotechnology Corporation with a rapidly passed special law. The nanotech project has been declared as being as important as the space project of the 1960s and has been allocated huge resources, comparable to the 2007 budget of the RAS.

The scale of the project makes it possible to create a new sector of basic and applied science outside the RAS. The nanotech corporation is a non-profit organization with unusually broad rights. It can finance any S&T project, establish new organizations and improve R&D teaching. In particular, it can support the development of research in universities, which have historically been underfunded. In my view, it is important for students to be taught and inspired by active researchers, and university research can increase the productivity of facilities.

A key characteristic of the new science policy is linking R&D to the demands of markets and society. Private-sector firms in the United States, Japan and Europe account for 75%, 71% and 65% of R&D spending, respectively[2]. In Russia the private sector accounts for around 30% of total funding — slightly less than it did in 1995. The public purse still accounts for 62% of research spending, whereas countries that are part of the Organisation for Economic Co-operation and Development average about 30%.

Not surprisingly, Russian production of patents per capita is 10 times lower than in Spain, and 60 times lower than in South Korea[3]. Despite having a similar proportion of researchers to Germany, German researchers produce 100 times as many patents per capita. The low level of Russian innovation is a hangover from weak international competitiveness during Soviet times, and the wide gap between much of R&D and the needs of the market.

Innovation is also being pursued by the Ministry of Economic Development and Trade. It has established a $600-million venture company — a state corporation for creating new private-venture funds — and five special 'tax-free' zones for S&T firms, including the former science cities of Zelenograd, Tomsk and Dubna. In 2006, the state created the Russian aircraft corporation (OAK), bringing dozens of existing plants, design bureaus and institutes under one umbrella organization. Moves to create similar corporations in shipbuilding, atomic energy production and other fields are underway.

Cooperation or monopolization?

My impression is that this new system of state capitalism in Russia is very similar to the South Korean 'chaebol' model — with government-assisted business conglomerates — of the 1990s. Generally speaking, it may help Russia to enter some global high-tech niche markets, especially if we focus on broad international cooperation, and forget the Soviet strategy of making everything ourselves. A successful example here is the new Superjet 100 commercial aircraft built by Sukhoi corporation (formerly the Sukhoi design bureau and now part of OAK) along with many international partners.

But the state domination of business and R&D may lead to monopolization and bureaucratization of S&T. We may even get a reprise of the Soviet administrative system. To avoid these negative consequences it is essential to encourage private investment in R&D and support for small innovative businesses. In basic science, the main tasks are structural reform of the RAS and creation of research universities. We need to identify and trust young leaders who have international experience, to solve these tasks. Only a new, young élite can help Russia become a modern knowledge economy. ■

Boris G. Saltykov is president of the Russian House for International Scientific and Technological Cooperation, Briusov per. 11, 125009 Moscow, Russia.

1. Science Indicators 2007 (State Univ., Higher School of Economics, Moscow, 2007).
2. Kouzminov, Y. *et al. The New Economy — A Chance for Russia.* (State Univ., Higher School of Economics, Moscow, 2003).
3. http://ns.worldbank.org.ru/files/rer/RER_13_eng.pdf

See Editorial, page 507.

©2007 Nature Publishing Group

BOOKS & ARTS

A celebration of Sputnik's fiftieth birthday

The launch of the first satellite sparked rejoicing worldwide but frayed some nerves in the West.

Red Moon Rising: Sputnik and the Hidden Rivalries that Ignited the Space Age
by Matthew Brzezinski
Times Books/Bloomsbury: 2007. 336 pp. $26/£17.99

William E. Burrows

The flight of the first Sputnik satellite in 1957, like the discoveries of electricity and nuclear energy, gave little hint of the leap it would start in civilization's development — except perhaps to a handful of rocket pioneers and science-fiction aficionados. Its successors would carry people to the Moon, orbit Earth as many hundreds of manned and unmanned mechanical servants, and conduct the otherwise unimaginable exploration of the Solar System.

One of the pioneers was Sputnik's creator, the articulate and infinitely imaginative Mikhail Tikhonravov. (To Matthew Brzezinski's credit, he dispels the old myth that Sergei Korolev was the mastermind behind Sputnik and lays that bouquet on Tikhonravov's grave, where it belongs. Tikhonravov's pioneering studies of orbital flight long predate Korolev's influence). Tikhonravov understood that the world changed forever on 4 October 1957, the International Geophysical Year. "This date," he said, "has become one of the most glorious in the history of humanity." Scientists around the world shared his enthusiasm: they knew that access to space would provide a cornucopia of understanding about our planet and the worlds around it.

Two days after the launch of 'primitive satellite number 1', as it was known to its operators, Washington newsman Eric Sevareid noted that the event had sparked a range of emotions: "Here in the capital responsible men think and talk of little but the metal spheroid that now looms larger in the eye of the mind than the planet it circles around. Men are divided in their feelings between those who rejoice and those who worry." He described scientists as mostly being "in raptures that the nascent, godlike instinct of *Homo sapiens* has driven him from his primordial mud to break, at last, the bound of his earth." Merton E. Davies, who was on Voyager 2's imaging team when the spacecraft did its 12-year grand tour of the outer planets, put it thus: "The joy of exploration is finding answers for which there are no questions."

As Brzezinski makes clear in his masterful history of the project, the glorious feat was mired in ignorance and petty bureaucratic politics at home. And it severely frayed nerves in the West, particularly in the United States. Many feared immediately that Sputnik was some sort of death-dealing weapon. Later they worried that the massive R-7 rocket that lifted it into space could just as easily carry a nuclear warhead due westward or over the Arctic. As it turns out, there is an uncanny similarity between R-7 and intercontinental ballistic missiles.

"The glorious feat was mired in ignorance and petty bureaucratic politics."

"Sputnik contained one final element that no ambitious newsman could resist: fear," explains Brzezinski, formerly *The Wall Street Journal*'s man in Moscow, "The missile that lofted Sputnik into space shattered America's sense of invulnerability". For the first time, the United States mainland lay exposed to enemy fire. In that respect, Russia's rockets were infinitely more frightening than the Japanese bombers that had attacked Pearl Harbor 16 years before. It was not distant naval bases on Pacific islands that were at risk, but the impregnable heartland itself: Cincinnati, St Louis, Chicago, Detroit, places that had never before had occasion to worry about foreign aggression.

The US Army had long since come to realize that the V-2 rockets that Germany rained down on the south of England near the end of the Second World War were actually very long-range artillery, the weapon of the future. In 1945, it took the precaution of grabbing Wernher von Braun and roughly 100 of his best engineers, together with a mountain of their V-2 missile engineering drawings. In 1945, in an operation code-named 'Paperclip' at Huntsville, Alabama, von Braun's team was hurriedly tasked to design the equivalent of the R-7, which they called Jupiter.

TIME & LIFE PICTURES/GETTY IMAGES

Sputnik's legacy goes far beyond being the first man-made machine in space, more than justifying the long-range historical perspective of *Red Moon Rising* (see page 542). Sputnik and its immediate successors were the formative ancestors of the aptly named spaceships that extended the human presence beyond the confines of the home planet. It was nothing less than the device that started the space age. That is cause indeed for a fiftieth birthday celebration. ■

William E. Burrows is in the Department of Journalism at New York University, 20 Cooper Square, New York, NY 10003, USA. He is founder and director emeritus of the university's Science, Health and Environmental Reporting Program and author of *This New Ocean: The Story of the First Space Age*.

©2007 Nature Publishing Group

... ALSO OUT NOW

A Ball, a Dog, and a Monkey: 1957 — The Space Race Begins
by Michael D'Antonio
Simon & Schuster

Epic Rivalry: The Inside Story of the Soviet and American Space Race
by Von Hardesty and Gene Eisman
National Geographic

Space: The First 50 Years
by Patrick Moore & H. J. P. Arnold
Mitchell Beazley

Von Braun: Dreamer of Space, Engineer of War
by Michael J. Neufeld
Knopf

After Sputnik: 50 Years of the Space Age
edited by Martin Collins
Smithsonian/HarperCollins

MITCHELL BEAZLEY & SMITHSONIAN/HARPERCOLLINS

FILM

Earth's fearful travelling companion

The Fever of '57
directed by David Hoffman
Available on DVD

In the Shadow of the Moon
directed by David Sington
On general release

Giovanni F. Bignami

"*Beep-Beep! Beep-Beep! Around the Earth it goes! Beep-Beep! Beep-Beep! And that's how science grows…*" So warbles the long-forgotten Sputnik song over the end credits of *The Fever of '57*. This rigorous and affecting documentary feature captures the US public's wonder — and then fear — on discovering that the Soviets had launched an 83-kilogram metal sphere over the world.

Named in poetic Russian style, *Sputnik zemli* means 'a travelling companion of the Earth'. Director David Hoffman uses rare archive footage, often from private collectors, to recount the drama of the launch in the context of the escalating cold war. He contrasts the awe of the man on the street for the majestic technological achievement with the political nightmare that began for the Eisenhower administration, already mired in civil-rights unrest.

Black-and-white images show worry etched into President Dwight Eisenhower's usually reassuring face. His embarrassment in being unable to answer press-conference questions calls to mind George W. Bush's demeanour on learning of the 9/11 attacks. Experts, including Edward Teller, immediately grasped the implications of getting cargo of this size into orbit: international ballistic missiles. The United States, the proclaimed science capital of the world, was humiliated.

Hoffman exhumes typical Khrushchev-era propaganda — tastefully shot, with a few too many Lenin portraits. Anti-American footage contrasts US prisons and hungry queues with food being tipped into the sea and the rich having their pets professionally manicured.

Political tensions heightened a month after Sputnik, when Laika the dog went into orbit in a full-bodied spaceship. US news coverage, now in colour, shows animal-lovers protesting.

Eisenhower was loath to seek help from the army, fearing that military involvement would stimulate an arms race. The US Army rocket programme was led by Wernher von Braun, the designer of Germany's V-2, which had bombed London and other cities 15 years earlier. Finally, von Braun's rocket launched the first US satellite, Explorer 1, on 31 January 1958. James Van Allen put Geiger counters in it and so discovered the radiation belts around Earth that now carry his name.

Soviet awe of Sputnik contrasted with US fear and political unrest.

VARIED DIRECTIONS INTL

Soon afterwards, Eisenhower created NASA as a civilian space agency. Reluctantly, von Braun joined, developing the Saturn V rocket that eventually brought man to the Moon. In an extraordinary public-relations coup, Eisenhower plotted the Scope mission with the nascent NASA in secret — not even Congress knew. The satellite bore his 1958 Christmas message to the world.

The US public was still worried. The Soviets were pushing farther into space. Both sides were developing nuclear weapons, as Eisenhower had predicted. During 1958, one nuclear test was conducted every three days, including the largest H-bomb ever detonated above ground. And both sides were developing civil defence programmes.

Meanwhile, headlines were screaming 'Soviets could claim the Moon'. In 1959 the USSR launched three Moon probes. One missed, another crash-landed on it (carrying a red flag, they say); the third sent us the first, breathtaking, image of the hidden side of the Moon.

David Sington's *In the Shadow of the Moon* picks up where *The Fever of '57* leaves off. This *cinéma-vérité* history of NASA's Apollo programme splices remastered original footage and original interviews with ten Apollo astronauts, presented by Ron Howard, director of the 1995 film *Apollo 13*.

As in the 1983 movie *The Right Stuff*, we learn what it took in the pioneering days of manned space missions to become a NASA astronaut. "An unshakable belief in your own infallibility," says one interviewee. The astronauts were all military pilots; more than one felt guilty about training for glamour and fame while their comrades were dying in the Vietnam War.

Notable for his absence is Neil Armstrong. Shown here is the eerie speech by President Richard Nixon, recorded for use in the event that the one-shot ascent module motor failed to ignite, leaving Armstrong and Buzz Aldrin to die on the Moon.

"We have yet to return to the Moon," says Sington at the end of his Sundance award-winning mosaic. Today's white manes and bald pates — yesterday's young men — are the only people to have hidden the whole Earth behind their thumb. NASA is working on "Apollo on steroids", with infinitely better technology than 40 years ago. The distance to the Moon, within a few centimetres, remains the same. ■

Giovanni F. Bignami is president of the Italian Space Agency and is at the Institute for Advance Study, 27100 Pavia, Italy.

©2007 Nature Publishing Group

The perpetual dance of continents

Supercontinent: Ten Billion Years in the Life of our Planet
by Ted Nield
Granta Books/Harvard University Press: 2007. 352/304 pp. £18.99/$29.95

David Oldroyd

As a geologist turned science journalist, editor and provocative blogger, Ted Nield has a complex view of life and science. His skills as a writer successfully convey in *Supercontinent* the recent exciting work in grand-scale geoscience to a wide scientific audience.

Nield's task is immense: to give an intelligible account of Earth's history, beyond the familiar story of wandering continents, drift and the simpler ideas of plate tectonics. He tells us what is thought to have happened way before Alfred Wegener's Pangaea (pictured here), following the trail of analogous earlier 'supercontinental' land masses: Pannotia, Rodinia, Atlantica, Columbia, Nena and Ur — the last probably appearing fairly soon after Earth's formation. To embrace a span of 10 billion years, Nield also envisages a whole succession of future supercontinents. Indeed, the next assemblage has already been named: Novopangaea.

Geoscientists have long been reconstructing continents, and their results are wonderfully displayed at various websites (such as www.scotese.com). Nield wants to help readers understand just how such reconstructions have been produced and how more recent work is extending reconstructions into the future as well as the past. He links the perpetual 'quadrille' of the continents to plate movements, climatic and atmospheric changes, and life's causal role in the changes. Everything is driven by the great convection currents within Earth's mantle.

Thus *Supercontinent* recounts geologists' debates about the 'introversion' and 'extroversion' of continents and describes the intricate geochemical and geomagnetic work that underpins their endeavours. Fearlessly, Nield plunges into the complex problems of the role of ophiolites, pieces of oceanic plate thrust onto the edge of continental plates, in revealing the histories of continents and oceans. He describes what he calls 'Rust World', when banded ironstone formations emerged 3.5 to 1.5 billion years ago, owing to the production of oxygen in the atmosphere by simple photosynthesizing life-forms.

There followed the age of stromatolites, dubbed here 'Lasagne World'. Then a period of cooling, when organisms sapped the atmosphere of the greenhouse gas carbon dioxide, and sunlight reaching Earth was reduced by dust from volcanic eruptions accompanying the break-up of Rodinia. So came 'Iceworld' or 'Snowball Earth'. This world is thought to have had several episodes during the period 710–580 million years ago, the climatic swings being due to the opposing effects of outbursts of volcanic gases and the sequestration of carbon dioxide by living organisms. The huge swings supposedly ceased when Rodinia fractured and landmasses were no longer mostly located near the Equator. The evolution of 'modern' life-forms quickly followed.

All this is admirably described. The attempted reconstructions of past and future continents and oceans is a major field of activity in contemporary geoscience. To handle it without oversimplification or getting lost in a maze of detail is no small accomplishment. I know of no other attempt to reduce the complexities of the relevant primary literature to the confines of a single popular-science book. However, one major element is glossed over: why do supercontinents form? They are thought to break up owing to radiogenic heat trapped beneath. But why aren't there perpetual random associations of land masses going on all over the globe? Why should there be a series of continental coalescences? The 'introversion' and 'extroversion' models partly address these questions, but undoubtedly there is much more to be said about them.

Nield rightly deems it important for the public to grasp the outlines of current (geo)scientific thinking and debate. Presumably to sweeten the pill, he brings key people into the story to good effect, incorporating contributions from the likes of Lord Kelvin, Wegener, John Joly, Tuzo Wilson, Janet Watson, John Sutton, Paul Hoffman and Christopher Scotese. The idea of supercontinents (or other continents) has also popped up elsewhere. So *Supercontinent* takes us on some entertaining — if, frankly, dead-end — trips to make this point. We are told of the thoughts of Madame Blavatsky, *l'abbé* de Bourbourg, Ignatius Donnelly and James Churchward about such notions as Lemuria, Atlantis and Mu. The strange *Urantia Book*, associated with a US cult, contains passages that by chance foreshadowed modern ideas on Rodinia. Such matters are not worth much attention, but I applaud Nield's brief swipe at American creationists.

Supercontinent warrants a special place in airport bookshops to stimulate travellers to contemplate the impermanence of the continents and seas they are flying over. ■

C. DARKIN/SPL

David Oldroyd is honorary visiting professor of history and philosophy of science at University of New South Wales, Sydney, New South Wales 2052, Australia, and author of *Thinking About the Earth: A History of Ideas in Geology* and *Earth Cycles: A Historical Perspective.*

©2007 Nature Publishing Group

SCIENCE IN CULTURE

A dog's life

Laika, the doomed stray, has achieved a kind of immortality.

Space agencies rely on the public's interest in people and animals to sustain engagement with their programmes. Curiosity over the ill-fated Mars lander Beagle 2, for example, was undoubtedly boosted by its clever name. Striking images of astronauts and space animals have strongly contributed to the visual output from the agencies.

On 3 November 1957, the fortieth anniversary of the Russian Revolution, a dog was launched into Earth orbit in Sputnik 2. Russia had already launched numerous dogs into suborbital flight, with at least four fatalities. To place a dog in orbit was an achievement of an altogether different dimension, especially only a month after the launch of Sputnik 1.

The dog nicknamed Laika ('Barker'), called 'Muttnik' in the American press, was selected from the cadre of trained dogs, all small female strays off the streets of Moscow. The BBC reported, "It is believed the Russians are planning to catapult the dog back to Earth, although there has been no official announcement confirming this".

Predictably, animal-welfare groups were outraged. In Britain, the National Canine Defence League called for a minute's silence for each of the dog's days in space. The RSPCA advised protesters to assemble at the Russian embassy.

Russian news releases, reluctantly acknowledging that Laika could not be returned to Earth, indicated that she remained healthy for some days. However, at the World Space Congress in Houston, Texas, in October 2002, Dimitri Malashenkov from the Institute for Biological Problems in Moscow admitted that Laika, stressed and overheated, probably survived for a maximum of seven hours.

The main photographs issued by the space agency show the attractive mongrel alert in her space harness and resting comfortably in her padded capsule. The former, which seemed to exude an air of bright courage, proved to be the more iconic. It was used on Mongolian and Romanian postage stamps, and to adorn souvenirs.

Heroically, Laika features prominently among the human space pioneers on the Monument to the Conquerors of Space in Moscow. Shortly after the launch of Sputnik 1, a competition was announced for the design of the monument. The winning entry was submitted by the sculptor A. P. Faidysh-Krandievsky and architects A. N. Kolchin and M. O. Barshch. The monument was inaugurated on 4 October 1964, the seventh anniversary of the Sputnik 1 launch.

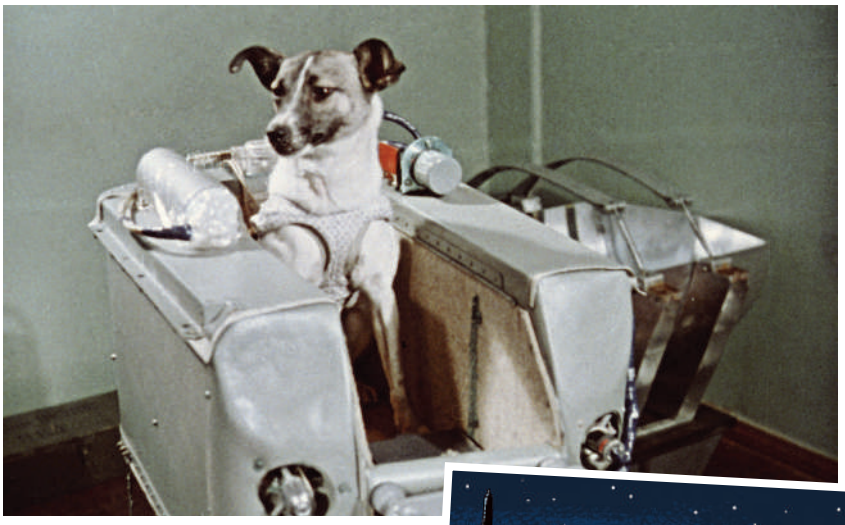
The monument assumes the form of a giant obelisk. At its summit is an ascending rocket, supported by a huge, tapering titanium slipstream. Around its base, which houses a museum, are sculpted reliefs of the space heroes. Laika's turned head and trace of the space harness are based on the classic photograph, but she is portrayed sitting within a simplified version of the capsule from the other image.

Laika has given her name to several pop groups, including the Finnish band Laika and the Cosmonauts and a successful UK-based ensemble founded in 1993 by Margaret Fiedler and Guy Fixsen. Even NASA named a soil target on Mars after her. Laika now even hosts a site on MySpace (http://profile.myspace.com/index.cfm?fuseaction=user.viewprofile&friendID=109345281).

Laika, the doomed stray, has achieved a kind of immortality. Many humans over the ages have traded their lives for great causes, particularly religious martyrs, confident in their afterlife and that they will live on through the renown of their sacrifice. For a conscious being, such a choice is an option.

We may be comforted by the 'portraits' of Laika, and perhaps moved by their pathos, just as we are when looking at images of revered human predecessors. There is a long tradition of memorializing animals, but such posthumous fame will not mean anything to a dog. ■

Martin Kemp is professor of the history of art at the University of Oxford, Oxford OX1 1PT, UK, and author of *The Human Animal in Art and Science*.



RIA NOVOSTI/AKG-IMAGES

N. BEER/CORBIS

FIRST SECOND/MACMILLAN BOOKS

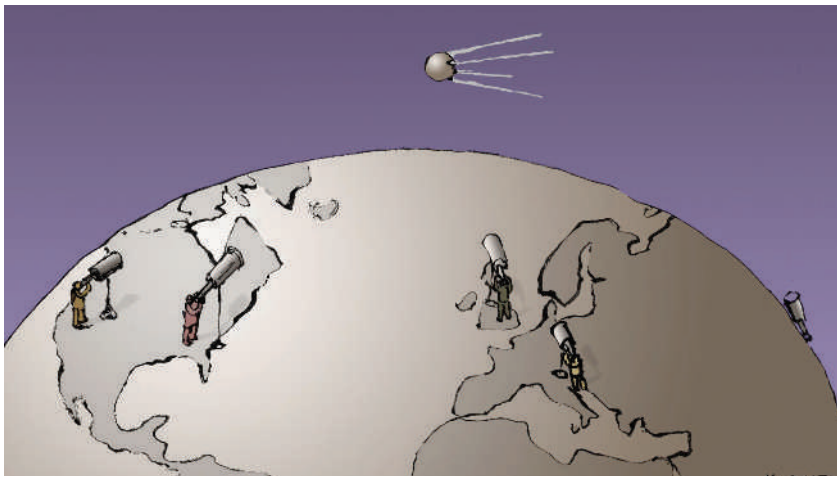
Laika lives on: in her padded capsule (top), on a monument to space heroes (left) and in an upcoming graphic novel.

©2007 Nature Publishing Group

The little ball made science bigger

When the Sputnik satellite went into orbit in 1957, it revolutionized the practice of international science and changed the demography of Western research.

Alexei Kojevnikov



J. KAPUSTA

The original Sputnik, a polished metal sphere with radio transmitters and antennas that orbited Earth, was a relatively humble venture. It was launched 50 years ago, partly by chance and before the space race, as publicly perceived, even existed.

The lead designers of the Soviet Union's R-7 intercontinental ballistic missile, including Sergei Korolev and Valentin Glushko, understood that their main mission was to develop strategic defence of the Soviet homeland, not cosmonautics. Capable of reaching the American continent, their machine would offer, for the first time, the possibility of retaliation against and deterrence of nuclear bombs from US forward airbases in Europe and Asia.

Some of the engineers still remembered their dream of space travel though, which at around 1930 had brought them as youngsters into the then amateur field of rocket design. At an opportune moment, when the Soviet leader Nikita Khrushchev was particularly pleased with progress, Korolev suggested using one of the missile's future tests for launching an artificial satellite into space.

Politicians demanded reassurances that such a distraction would not delay the main job, but agreed to reward the scientists and engineers — even though their desire seemed somewhat childish. Without waiting for more sophisticated instruments to be developed, Korolev went ahead with launching what was formally referred to as the 'simplest sputnik' — a basketball-sized satellite that could confirm the possibility of radio communication through the ionosphere. Although the R-7 was at a stage when roughly every second test encountered problems, the launch of Sputnik on 4 October 1957 went smoothly.

In the eyes of the Western public, Sputnik was a very different story. Even the Soviet engineers, aware that they had accomplished something very important, did not anticipate the enormity of the political reaction. Typically, the Soviets had looked down on the Western media's propensity for sensationalism. In this case, they found it working in their favour and started supporting it with propaganda claiming that the socialist system was taking the lead in technological development. Later, as the public and media attention shifted towards manned flight, the importance of Sputnik was unjustly downplayed.

The cold-war mentality is partly responsible for Russia and the United States choosing to celebrate the 1961 orbital flight by Yuri Gagarin and the 1969 moonwalk by Neil Armstrong as their respective victories in the space race. Instead, both should have recognized Sputnik as humanity's first great breakthrough in space. In the past 50 years, the series of sputnik missions has proved invaluable — from allowing global communication to fostering our common environmental awareness. The purpose of manned missions has remained uncertain.

Like the discovery of DNA's structure, Sputnik changed the practice of international science. Although a product of military research, it initiated a period of relative decline in the militarization of science that had become so overwhelming in the wake of the 1945 atomic bomb. The contrast between the indifference with which the public greeted the announcement of R-7's first successful military test, and the excitement generated two months later when the same missile launched Sputnik, spurred politicians to direct more support towards science through at least nominally civilian agencies. R-7 was eventually found to be impractical for military purposes and was replaced by other designs, but it remains the most successful, reliable and frequently used vehicle in the history of space launches.

"Many of the scientists now retiring owe their jobs to the launch."

Some science policy concepts that were previously thought specific to Soviet culture were borrowed internationally, first by the United States. For decades the Soviets had been accused of denying the value of 'pure research'. Until about 1960 they were ideologically committed to the notion that basic research is inseparable from the development of technology and other useful applications. In the wake of the Sputnik launch, the idea of separating pure and applied science began to lose ground among science policy-makers in the West. The notion that the two kinds of research should be considered and pursued separately has all but disappeared in the current era of genomics and nanotechnology.

Sputnik also changed the demography of the scientific profession. The Soviet success was attributed to a scientific education system that had for some time been producing more highly trained scientists and engineers per capita than the West. The communist belief that science should be treated as a mass profession, not an elitist activity, led to wide recruitment from lower classes, women and ethnic minorities.

A corresponding shift in the United States came partly from the National Defense Education Act of 1958, which aimed to boost education at all levels. More importantly, changes in immigration laws led to an influx of trained professionals from Europe, Asia and later other regions.

A big portion of the generation of scientists now retiring, and those who are replacing them, owe their jobs to the launch of the little aluminium sphere in 1957. That science in the West is now much larger, and more multiracial and multicultural, is the most important — if unintended — consequence of Sputnik. ■

Alexei Kojevnikov is at the University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada.

See also pages 536 and 538.

ESSAY

©2007 Nature Publishing Group

How the mighty have fallen

In 1957, science advisers were brought into the White House as the President's Science Advisory Committee. Its demise has deprived the US government of invaluable counsel.

Richard Garwin

The Second World War left the United States with enormous energy among its scientists and great advances in technology — for example, radar, the proximity fuse, nuclear weapons and the ability to detect sound underwater. Taking office in 1953, President Dwight Eisenhower consulted the little-known Science Advisory Committee (SAC) in the Office of Defense Mobilization. Eisenhower depended on 'his scientists' to help him counter those military officers and hard-sell technologists whose enthusiasm to protect the country and to advance the state of technology often exceeded their competence.

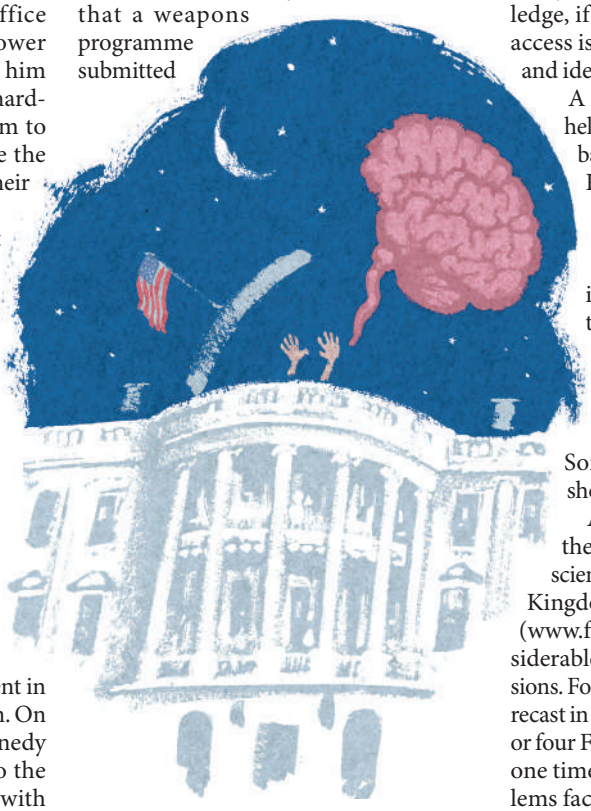
After the launch of the Soviet satellite Sputnik on 4 October 1957, Eisenhower brought the SAC into the White House as the President's Science Advisory Committee (PSAC). The PSAC had a key role in creating NASA in 1958. From then until 1973, when it was disbanded by President Richard Nixon, it helped to lay out a civil scientific space programme and to guide military space intelligence.

Although often dismissed by critics as the 'scientists' lobby in the White House', the PSAC concentrated on marshalling science for government, as seen in the committee's involvement in the first manned mission to the Moon. On 25 May 1961, President John F. Kennedy called for a mission to send man to the Moon within the decade. In a deal with his science adviser Jerome Wiesner, chair of the PSAC, Kennedy never claimed that the purpose of the Apollo programme was to advance science, and the PSAC never criticized it publicly despite internal reservations.

One of the committee's earliest achievements was to place scientific expertise at the highest level in the US Department of Defense in the form of a director of defence research and engineering. This post was a powerful one — number three in the department — and the PSAC ensured that its first occupant was the young nuclear physicist Herbert York, followed by fellow physicists Harold Brown and John Foster.

Opposition to the PSAC arose early on, in the form of complaints to members of Congress and government officials from scientists and industry adversely affected by PSAC analyses. Paradoxically, the higher profile of science and technology that the PSAC helped to create in government departments and agencies made it easier for its opponents to argue that the White House didn't need its own scientific advisers.

In fact, such advice was never more needed. Towards the end of the 1970s, I was told by one of our most competent defence secretaries that he did not regard it as his responsibility to ensure that a weapons programme submitted by the Department of Defense for the president's budget was sound and worthwhile. He saw it as the job of the White House Office of Management and Budget to sort that out, with the help of the PSAC. To me this was evidence that increased scientific and technological expertise in government departments increased rather than reduced the need for scientific competence in the White House.



D. PARKINS

Since Nixon abolished the PSAC, no president has seen fit to create an advisory committee of a similar level of commitment or energy — despite the establishment of a President's Council of Advisors on Science and Technology by George Bush Snr in 1990. In 1972, Congress had founded the Office of Technology Assessment (OTA), which was run by a bipartisan congressional board. Its studies were conducted by a small staff supplemented by people hired for individual projects. The unique feature of the OTA was an advisory committee for each study that included passionate proponents and opponents of the programme in question, whose points of view were evaluated, if not endorsed.

The United States has drawn particular strength from granting independent technical consultants access to government at many levels. Such individuals add knowledge, if not always wisdom. Today, that access is increasingly subject to political and ideological tests.

A strong PSAC would have been helpful in the controversy over space-based missile defences that followed President Ronald Reagan's 'Star Wars' speech of 23 March 1983 (although there were good OTA studies on this) and more recently in setting a response to potential threats of bioterrorism and disease pandemics.

In a further blow, the OTA was abolished in 1995 by a Congress under Newt Gingrich's leadership. Some likened this act to Congress shooting themselves in the brain.

Against this backdrop of decay in the United States, it is good to see a scientific advisory system in the United Kingdom — the Foresight programme (www.foresight.gov.uk) — that has considerable influence on government decisions. Foresight was established in 1978 and recast in its present form in 2002. The three or four Foresight projects under way at any one time address serious potential problems facing the country. Curent projects include mental health and wellbeing, obesity, and sustainable energy management. A sponsoring government minister for each project ensures that the advice gets to people with the power to act on it.

Foresight stands in sharp contrast to the absence of analysis in the United States for decision-making in and direction of programmes such as missile defence, smallpox-vaccine production and homeland security. The demise of integrity and competence in the US government is a matter for dismay. The restoration of the OTA would be a partial remedy — and a harbinger of reform. ■

Richard Garwin is IBM fellow emeritus at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA and a former member of the PSAC.

For more essays and information see http://nature.com/nature/focus/arts/scipol/index.html

©2007 Nature Publishing Group

NEWS & VIEWS

Foot and mouth — an argument for analgesics.

HULTON-DEUTSCH COLLECTION/CORBIS

NEUROSCIENCE

A local route to pain relief

Edwin W. McCleskey

Local anaesthetics stop pain, but block all other sensations too. In rats, one molecular delivery vehicle makes an unusual local anaesthetic specific for pain — provided a little spice is added to the mix first.

Medicine has no shortage of great anaesthetics: they have been making surgery tolerable by eliminating consciousness, or by blocking complete nerve systems, for 160 years. But what we do need more of are good analgesics: drugs that suppress pain without affecting any other sensation. On page 607 of this issue, Binshtok and his colleagues — based at the Massachusetts General Hospital in Boston, where surgical anaesthesia was first demonstrated in 1846 — report a new approach to analgesia[1]. They have found a way to deliver a local anaesthetic so that it blocks pain sensing alone.

To those in the know, an anaesthetic seems the last place to look for an analgesic effect. When you are numbed for the dentist's drill, the anaesthetic used blocks all the voltage-gated sodium channels in a nerve. These are proteins that conduct action potentials in nerve axons by timing, with submillisecond precision, the flow of current in the form of sodium ions across cell membranes. Blocking all sodium channels blocks all sensation; it blocks movement if the affected nerve includes motor axons, or the responses of internal organs if a controlling 'autonomic' nerve is exposed. Recently, however, certain sodium channels have been identified as being present only in pain-sensing neurons[2]. Drug companies have thus begun to comb their libraries of uncommercialized local anaesthetics for a 'magic bullet' drug that affects pain alone.

Binshtok *et al.*[1] follow a different path. They study the TRPV1 ion channel, which is opened up by capsaicin — the chemical that makes chilli peppers spicy and creates the burning sensation when they are rubbed on skin — or by temperatures that rise above 42 °C, the skin temperature that we consider unpleasantly warm. The temperature-dependent gating of TRPV1 and its specific expression profile, as demonstrated in a series of animal and human studies[3], show that it is a molecular sensor for noxious heat that is expressed only on small nociceptors. Nociceptors are sensory neurons that translate noxious stimuli into action potentials and conduct these electrical signals from the site of stimulation to the spinal cord. Small nociceptors conduct slowly (at about 1 m s^{-1}), and TRPV1-positive neurons therefore mediate slowly developing, persistent pain — anything from the half-minute of smarting after stubbing a toe to the never-ending discomfort of an arthritic joint.

Crucial to the story is that, although the pore of the TRPV1 channel rejects negatively charged anions, it promiscuously allows most positively charged inorganic cations to pass — and even organic cations as large as some local anaesthetics. In most cases, this is irrelevant: local anaesthetics are generally weak acids that occur in both a positively charged (protonated) form and an uncharged, 'free-base' form. As free bases, local anaesthetics readily permeate lipid cell membranes without the need to pass through the pores of an open ion channel. This is why they shut down sodium channels in all neurons; indeed, the potency of an anaesthetic rises in proportion to its lipid solubility.

Once an anaesthetic has passed through the membrane, it blocks the sodium channel from within by docking in a wide hydrophobic vestibule just to the inner side of the sodium selectivity filter, the narrowest part of the channel's pore[4]. This inner vestibule, whose existence was first deduced in potassium channels[5] and has now been seen in their crystal structure[6], is common to all voltage-gated ion channels and is the binding site for many different drugs used against various voltage-gated channels[7].

But there is one particular anaesthetic, known as QX-314, that cannot infiltrate lipid membranes under its own steam: it has a permanent positive charge, making it lipid-insoluble. As a result, it fails to block sodium

©2007 Nature Publishing Group

NATURE|Vol 449|4 October 2007

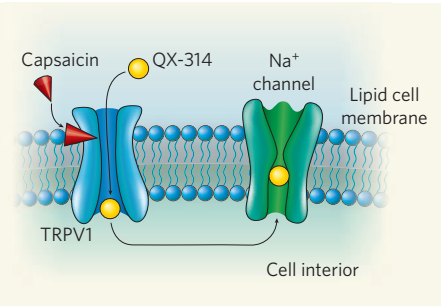


Figure 1 | A spicy solution. QX-314, a charged local anaesthetic, cannot pass through a lipid cell membrane. Instead, it gains access to the cell's interior through TRPV1, an ion channel that is opened by capsaicin, the 'hot' component in chilli peppers. Once inside, QX-314 can bind to the inner vestibule of sodium (Na^+) channels, the site for local anaesthetic action. This transmembrane access route for QX-314 exists only on nociceptors, pain-sensing neurons on which TRPV1 is expressed. Thus, as Binshtok *et al.*[1] show, administering a mix of QX-314 and capsaicin acts as an analgesic, shutting off pain sensation without affecting any other nerve responses.

channels when applied outside cells, although it succeeds when injected into them. Once inside, QX-314 binds within the hydrophobic vestibule just like any other anaesthetic. In fact, kinetic features of QX-314's blocking mechanism — dependence on the direction of current flow and on membrane voltage — were the initial evidence for the vestibule binding site of anaesthetics in general.

What if QX-314 could be delivered to the insides of cells through the large, cation-selective pores of the TRPV1 channel? By exploiting the fact that TRPV1 resides only on nociceptors, this might inhibit pain without other side effects. And Binshtok *et al.*[1] found that, although neither capsaicin nor QX-314 alone could shut off the sodium channel or action potentials, applied simultaneously to nociceptors they could do just that (Fig. 1). When injected into the foot or perfused onto a nerve of a rat, the mixture inhibited the animal's sensitivity to noxious thermal and mechanical stimuli without causing paralysis. This analgesia lasted about 2 hours.

The significance of this discovery might go beyond just the blockade of sodium channels. First, the same sort of strategy might prove useful for delivering other drugs to other voltage-gated channels. Second, local anaesthetics are more than simple ion-channel blockers. At low concentrations, they exhibit a clinically important selectivity, known as 'use-dependent inhibition', by binding with greater affinity to inactivated sodium channels than to those poised to open[8]. Low concentrations of anaesthetic thus selectively inhibit sodium channels in cells that are frequently firing action potentials, because those channels cycle more frequently through the inactivated state.

Such selective targeting of hyperactive cells explains why the local anaesthetic lidocaine, when perfused into the entire body at very low concentrations, suppresses certain cardiac arrhythmias without affecting nerves. Pathological pain — pain that is persistent but not caused by an existing injury — is another example of hyperactivity, this time in nociceptors. Nerve blocking by local anaesthetics is sometimes used to treat pathological pain because relief from the pain persists long after the initial numbness wears off; this might be because low, residual levels of local anaesthetic selectively inhibit nociceptors that are too active[9]. If Binshtok and colleagues[1] are right, the humble local anaesthetic may thus prove to have four mechanisms for specificity: through infusion onto a particular nerve for its most basic application; through use-dependence, as a remedy for electrical hyperactivity; through the targeting of distinct subtypes of sodium channel to alter the activity of distinct cells that express them; and now through the TRPV1 channel, a nociceptive drug-delivery vehicle that yields an analgesic effect.

But before we get carried away, the TRPV1 trick must first be shown to work in humans. In addition, the ideal cocktail of capsaicin and TRPV1 modulators must be found, in order to avoid any damage that capsaicin would cause at excessive levels. It will be useful, too, to optimize the use-dependence and TRPV1-permeability of QX-314. This drug, which was until now just an exotic reagent used by ion-channel biologists, will be the focus of a new effort in the search for better analgesics. ■

Edwin W. McCleskey is at the Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA.
e-mail: mccleskeye@hhmi.org

1. Binshtok, A. M., Bean, B. P. & Woolf, C. J. *Nature* **449,** 607–610 (2007).
2. Cummins, T. R., Sheets, P. L. & Waxman, S. G. *Pain* **131,** 243–257 (2007).
3. Caterina, M. J. & Julius, D. *Annu. Rev. Neurosci.* **24,** 487–517 (2001).
4. Strichartz, G. R. *J. Gen. Physiol.* **62,** 37–57 (1973).
5. Armstrong, C. M. *Q. Rev. Biophys.* **7,** 179–210 (1975).
6. Doyle, D. A. *et al. Science* **280,** 69–77 (1998).
7. Yu, F. H., Yarov-Yarovoy, V., Gutman, G. A. & Catterall, W. A. *Pharmacol. Rev.* **57,** 387–395 (2005).
8. Hille, B. *J. Gen. Physiol.* **69,** 497–515 (1977).
9. Yanagidate, F. & Strichartz, G. R. *Handb. Exp. Pharmacol.* **177,** 95–127 (2007).

ATOMIC PHYSICS

Cold meeting at a junction

Charles A. Sackett

The Josephson effect is a macroscopic manifestation of quantum mechanics usually seen in superconductors. Observation of this effect in a gas of ultracold atoms demonstrates the underlying unity of solid and gaseous systems.

On page 579 of this issue[1], Levy *et al.* report the observation of the Josephson effect in a cold atomic gas. Brian Josephson discovered the original version of this effect in 1962, when he was a young graduate student at the University of Cambridge[2], and it earned him a share of the 1973 Nobel Prize in Physics.

Josephson considered what happens when two superconducting plates are placed next to each other with an insulating layer between them. In a world determined by classical physics, this makes a simple capacitor that stores up static electric charges. But quantum mechanics warns us that it is hard to pin a particle down in any one place. Because of this, if the Josephson junction is formed with a sufficiently thin insulator, electrons from one plate can tunnel through the barrier to the other plate, resulting in a flow of current.

For plates made of normally conducting metal, this tunnelling is haphazard, and the effect is equivalent to that of a resistor shorting the capacitor. But in a superconductor, where current can flow without resistance, the quantum state of the electrons is highly correlated, and the tunnelling becomes coherent. Josephson's breakthrough was to realize that this meant interference could be observed, because the tunnelling wavefunction from one electrode combines with that from the other in a way that depends on their relative phase. This interference gives rise to two main effects. First, a steady current can flow through the junction even when no voltage is applied. Second, when a steady voltage is applied, an oscillating current results. These are known respectively as the d.c. and a.c. Josephson effects. They are at the heart of many important technologies, particularly in the measurement of electric and magnetic fields.

Levy *et al.*[1] do not use superconductors, but start with a gas of rubidium atoms in the form of a Bose–Einstein condensate. Condensates are a kind of superfluid, in which the atoms share a quantum wavefunction just as electrons do in a superconductor. The atoms are at a temperature just a few billionths of a degree above absolute zero, and are held in vacuum using a magnetic trap (a box to keep them from drifting away). Levy *et al.* divided this trap in two by sending a tightly focused laser beam through its centre. The beam formed a barrier for the atoms that was analogous to the insulating layer in the original Josephson effect; it was similarly narrow enough that atoms on one side had a non-zero probability of tunnelling through to the other.

Following a suggestion made by Giovanazzi

©2007 Nature Publishing Group

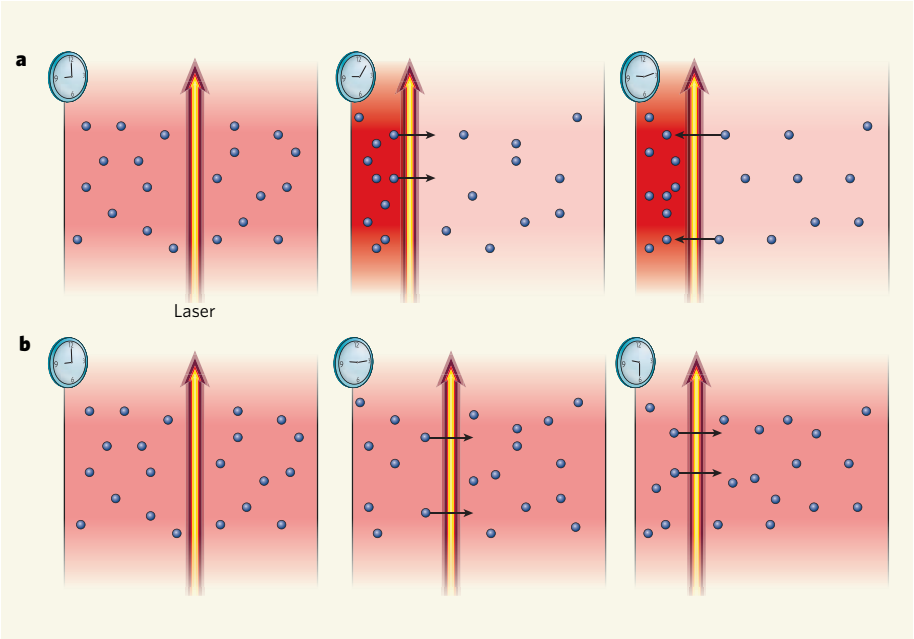


Figure 1 | Flow or no flow? In Levy and colleagues' experiments[1], a collection of ultracold atoms is held in a magnetic trap, and divided in two by a laser beam. **a**, If the atoms on one side are compressed quickly by displacing the beam, they can tunnel through the barrier to relieve the pressure. But owing to interference between the quantum wavefunctions of the atoms on either side of the beam, the flow later reverses itself, and then continues to oscillate back and forth without equilibrium ever being restored. This is the atomic-gas equivalent of the a.c. Josephson effect at a superconductor–insulator–superconductor junction. **b**, When the laser beam is moved more slowly, by contrast, atoms tunnel through at a steady rate, and the density remains the same throughout the gas. This flow occurs even though no pressure ever builds up to drive it — a 'supercurrent' analogous to the d.c. Josephson effect.

et al.[3], the authors observed the a.c. Josephson effect by moving the beam from the centre of the box towards one side (Fig. 1a). Atoms in a condensate normally repel one another, so the energy of the atoms on the side that was compressed increased as they were crowded against their neighbours. This interaction energy, also known as the chemical potential, served as the equivalent of the voltage in an electrical circuit: it made the compressed atoms want to flow across the barrier to the other side. But when the authors monitored the number of atoms on the compressed side, they saw that it did not simply decrease; instead, it oscillated in time. This is because the tunnelling rate depends on the interference of the two quantum waves representing the two separated condensates, and the sign of this interference oscillates in time when the condensates have different energies. Previous experiments had exposed the same sort of effect in an array of many condensates[4], but this is the first time that a simple, two-state system directly analogous to a superconducting Josephson junction has been observed.

Levy and colleagues' atomic junction[1] also exhibited the d.c. Josephson effect. In this case, the laser was again displaced to one side, but much more slowly (Fig. 1b). As it moved, the authors found that atoms were flowing through the barrier, but that the densities of the atoms, and therefore the chemical potentials of the two condensates, remained precisely equal. This is the analogue of a supercurrent — a flow without a voltage to drive it. They observed this behaviour as long as the beam was moved more slowly than a critical velocity of about 40 μm s^{-1}. A corresponding critical value of the current occurs in the superconducting case.

Even compared with superconductors, Bose–Einstein condensates are delicate and difficult to produce. So it is unlikely that the atomic Josephson effect will find applications as widespread as the electrical version. But the authors do point out the possibility of using the atomic effect to create an exquisitely sensitive rotation detector that might find use in 'inertial guidance' systems for rockets and aircraft.

Above all, however, this work represents a milestone in the effort to map the physics of solid matter onto atomic systems of ultracold gases. The advantage of this mapping is that atomic systems are much more easily controlled. For instance, it was a simple matter for Levy *et al.* to vary the height and width of their tunnelling barrier by adjusting the optics of their laser beam. In a superconducting system, such tests would require the fabrication of entirely new samples. More generally, laser beams can be used to create defect-free 'atomic crystals' of arbitrary symmetry and adjustable binding strength[5]. Even the interactions between atoms — analogous to the charge for electrons — can be tuned[6].

The hope is that a 'proving ground' might be created, where theories could be tested under various well-controlled conditions and where the complicated phenomena of real material systems could be sorted out. Levy *et al.*[1] have

50 YEARS AGO

Documentation in Action — "The world of thought (scholarship) and the world of action have always been interdependent", and this means that to provide for an efficient human society we require to build up an efficient scheme of communications. At the present time with the rapid increase of knowledge, and even more rapid increase of the volume of publications, this need is becoming even more imperative... The newly emerging science, dealing with the organization of human knowledge, which has not as yet received a distinctive name, includes a very large number of departmental aspects, such as...principles of classification, library organization, card systems and recording and calculating machines... The choice of terms used in this book ('processing', 'project', 'language engineering' and the verb 'to programme') brings us into close contact with the American vitality of language and action; but this need not deter the scientist, or engineer, or anyone dealing with organized knowledge or administration from studying this book.

From *Nature* 5 October 1957.

100 YEARS AGO

'The Fourteenth International Congress of Hygiene and Demography' — Under the heading of dietetics... the alcohol question also was to the fore, but a somewhat guarded opinion seemed to be expressed regarding its supposed evil effects when used in moderation. Dr. Triboulet, of Paris, considered that alcoholic drinks had an unfavourable action in cases of tuberculosis, and that alcohol diminishes resistance and predisposes to tuberculosis. Dr. Moeli, of Berlin, considered that alcohol is not a food, is not necessary, and in many cases is detrimental, and that its use should be discouraged in every way, not only by teaching the masses the evil effects of its use, but by instituting other interests so that the lack of it should not be felt.

From *Nature* 3 October 1907.

50 & 100 YEARS AGO

©2007 Nature Publishing Group

strengthened this link by showing that atomic condensates can mimic superconductors in one of their hallmark effects. ■

Charles A. Sackett is in the Department of Physics, University of Virginia, 382 McCormick Road, PO Box 400714, Charlottesville, Virginia 22904-4714, USA.
e-mail: sackett@virginia.edu

1. Levy, S., Lahoud, E., Shomroni, I. & Steinhauer, J. *Nature* **449,** 579–583 (2007).
2. Josephson, B. D. *Phys. Lett.* **1,** 251–253 (1962).
3. Giovanazzi, S., Smerzi, A. & Fantoni, S. *Phys. Rev. Lett.* **84,** 4521–4524 (2000).
4. Anderson, B. P. & Kasevich, M. A. *Science* **282,** 1686–1689 (1998).
5. Bloch, I. *Phys. World* **17,** 25–29 (2004).
6. Inouye, S. *et al. Nature* **392,** 151–154 (1998).

DEVELOPMENT

Inflationary pressures

David Bryant and Keith Mostov

Blowing up a balloon seems straightforward: pump in gas and let the changing air pressure do the rest. But when it comes to blowing up nature's own balloons — lung airways — things are a little more complicated.

The lung consists of an elaborate network of branching tubes; the tubes themselves are made of a layer of epithelial cells wrapped around a hollow, air-filled core, much like the thin rubber of a balloon surrounds the air inside. For the tubes to fit together, and for the right amount of gas to pass through them, they must have precisely the correct dimensions. Studies of developing embryos have revealed several molecular players that contribute to the construction of the lung-tube network[1]. Now, in a paper published in *Developmental Cell*, Tsarouhas *et al.*[2] tackle the long-standing question of how the tubes expand, and provide breathtaking movies of this dynamic process *in vivo*.

In biological tubes, epithelial cells sit in a uniform pattern. The top of the cells (the apical surface) is in contact with the tube core (the lumen) — and, in the case of lung tubes, with the air taken in — whereas the base of the cells abuts underlying blood vessels[3]. While the tubes are developing, their lumen is very narrow, with the apical surfaces of opposing cells close together. As the tubes mature, the lumen diameter increases to the required size. Until now, it has been largely unclear how this expansion occurs and which molecules are involved in its regulation.

Filming lung-tube expansion in developing fruitfly embryos, Tsarouhas *et al.*[2] provide a striking visual demonstration that this process involves three main, rapid events in a precise timeframe (Fig. 1). First, around 10.5 hours after eggs are laid, epithelial cells lining the tube spit out large amounts of protein into the tube lumen. Next, over the following 30 minutes, the tubes' diameter expands 2.5–3-fold. Finally, some 7.5 hours later, much of the material in the tube lumen is cleared out by the same epithelial cells. Shortly before the embryo is ready to hatch, the liquid remaining in the lung is exchanged for gas.

To identify the molecular regulators of lung-tube expansion, the authors studied mutant flies in which the expansion of the tubes was compromised. Their search converged on several molecules involved in an evolutionarily conserved biological process, biosynthetic vesicle formation, which mediates intracellular transport. Proteins are synthesized inside cells, and before they can be transported to their final destination (such as the lung-tube lumen) they must be loaded in membrane-enclosed vesicles. Such vesicles form by budding off from the membrane of the Golgi complex — an intracellular packaging station analogous to a central post office.

The authors find that when the synthesis of proteins that regulate the budding of a particular class of vesicle is perturbed, or mutant proteins are synthesized, the transport of proteins destined for the lumen is reduced, and tube expansion is disrupted. These findings highlight a previously unknown function of the vesicular transport machinery in tube luminal expansion.

Protein-filled tubes, however, are of little use as airways, and so the authors studied the regulators of the subsequent protein clearance. Endocytosis, a process in which a small portion of the cell membrane is sucked in to form a pouch and then pinched off to create an internal vesicle, is a central mechanism by which cells take up liquid and protein from their environment[4]. Endocytosis of proteins is also akin to organizing trash: what can be recycled is recycled, and what can't is sent to disposal compartments in the cell for degradation.

Tsarouhas *et al.* show that when the synthesis of many proteins involved in endocytosis is disrupted, or mutant proteins are synthesized, the clearance of material from tube lumina is largely abolished, and the tubes fail to mature into airways. Thus, the epithelial cells lining the tubes seem to have a more active role in tube expansion than was previously appreciated, as they coordinate both mass release of proteins to expand the developing tube lumen and subsequent 'swallowing' of this material to clear the mature lumen.

Static images of tracheal tube development have previously allowed many insights into lung development[1]. But the power of these latest findings[2] lies in the fact that they combined live imaging with the genetic malleability and transparent development of fruitfly embryos. The films they have recorded are a testament to the idea that 'seeing is

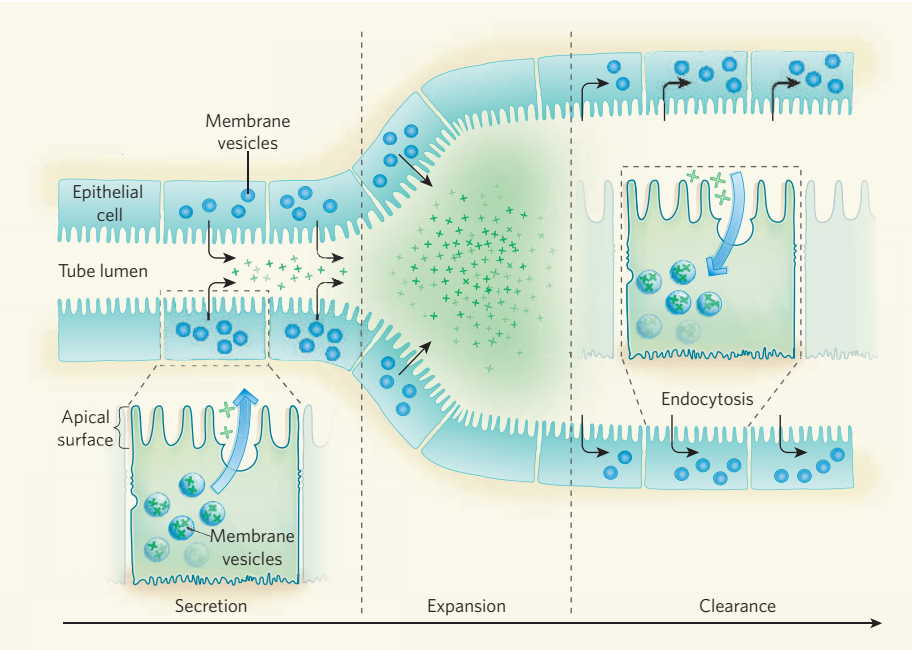


Figure 1 | Expanding airways. Tsarouhas *et al.*[2] show that during fruitfly embryonic development, the trachea matures in three stages. First, lining epithelial cells secrete large amounts of protein into the intra-luminal area of rudimentary tubes. Shortly after, the lumen diameter expands to its mature size. The protein mass is then internalized from the tube lumen by the process of endocytosis to make the airways ready to be filled with gas.

©2007 Nature Publishing Group

believing' — and, indeed, understanding — as they reveal many aspects of the airway-maturation process *in vivo*, an achievement that remains technically challenging in mammals.

Lumen expansion, especially in vertebrates, is probably controlled by various mechanisms. For example, chloride-ion secretion is involved in determining the diameter of the lumen in the thyroid gland[5], but so far there is little evidence for mass release of proteins into the lumen of most types of vertebrate tube. Expansion of the mammalian lung and the intestinal lumen in zebrafish[6,7] is apparently controlled by the movement of fluid into the lumen. So it is possible that the role of protein secretion, as seen by Tsarouhas *et al.*, is to deliver protein regulators of osmotic pressure to the apical surface of epithelial cells, which in turn allow the tubes to expand by filling with liquid.

In terrestrial animals, from flies to humans, air-carrying tubes must be cleared of liquid as the embryo transfers from the liquid environment of the egg, uterus or similar developmental domain, to breathing air. Mammalian lungs accomplish this by very rapid fluid absorption at birth[6]. Little protein needs to be removed from normal lungs, but after acute lung injury large amounts of protein accumulate and must be cleared from the lumen of the lung airways to restore their function. This may involve vigorous endocytosis by epithelial cells that is similar to the process described by Tsarouhas and colleagues.

Their work, along with previous studies, indicates that nature has given living organisms several ways to blow up their 'balloons'. But despite this, both across species and in different organs of the same organism, at least one principle remains the same — the need to reshape and remodel tubes to suit changing organ physiology. ■

David Bryant and Keith Mostov are in the Departments of Anatomy, and Biochemistry and Biophysics, University of California, San Francisco, California 94143-2140, USA.
e-mail: keith.mostov@ucsf.edu

1. Lubarsky, B. & Krasnow, M. A. *Cell* **112,** 19–28 (2003).
2. Tsarouhas, V. *et al. Dev. Cell* **13,** 214–225 (2007).
3. O'Brien, L. E., Zegers, M. M. P. & Mostov, K. E. *Nature Rev. Mol. Cell Biol.* **3,** 531–537 (2002).
4. Conner, S. D. & Schmid, S. L. *Nature* **422,** 37–44 (2003).
5. Yap, A. S., Stevenson, B. R., Armstrong, J. W., Keast, J. R. & Manley, S. W. *Exp. Cell Res.* **213,** 319–326 (1994).
6. Olver, R. E., Walters, D. V. & Wilson, M. S. *Annu. Rev. Physiol.* **66,** 77–101 (2004).
7. Bagnat, M., Cheung, I. D., Mostov, K. E. & Stainier, D. Y. R. *Nature Cell Biol.* **9,** 954–960 (2007).

MATERIALS SCIENCE

Colloidal crystals find new order

F. Schüth and F. Marlow

A deft colloidal templating process allows simple-cubic crystals to be formed from more readily available complex precursors. It's a promising way to produce the regular crystals much in demand for photonics.

Colloidal crystals — arrangements of generally spherical particles between around 10 and 100 nanometres in size — typically assume a dense face-centred-cubic (f.c.c.) order, with particles at each vertex and in the middle of each side of the crystal's constituent cubic cells. This kind of packing is tight, filling about 74% of the available space; but the small gaps left can be infiltrated by other, smaller particles to form a replica of the original f.c.c. template after this has been removed.

Writing in *Angewandte Chemie International Edition*, Li *et al.*[1] show how such a structure can be made to reassemble when its template is removed thermally (calcinated), to form a less densely packed, simple-cubic structure (Fig. 1). Their work represents a new way to synthesize crystals with structures that deviate from dense sphere packings[2], and that until now were hard to make. Such crystals are much in demand, for instance as photonic crystals that act on light as semiconductors do on electrons.

Binary nanoparticle systems of various compositions have already been assembled into different packing motifs[3,4]. But to create non-close-packed structures of larger particles (in the 100-nanometre range), either the forces in the system must be maintained in a delicate balance[5], or prestructured surfaces are required to act as a template[6].

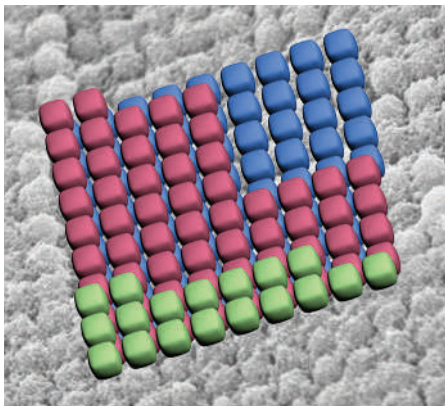


Figure 1 | Simple cubes. A scanning-electron-microscope image of Li and colleagues' simple-cubic packing[1] (different layers are emphasized in false colours), formed in the voids of a face-centred-cubic crystal removed by calcination.

A. STEIN

Li and colleagues' approach is different, and involves first infiltrating the pores of a close-packed colloidal crystal of poly(methylmethacrylate) (PMMA) spheres with a cocktail of chemicals in solution — titanium isopropoxide, triethylphosphate, the surfactant Brij 56 and acetylacetone. On calcination, this precursor is converted to a binary titanium dioxide–phosphorus pentoxide (TiO_2–P_2O_5) structure that assumes cuboid and spheroidal shapes in the two differently sized gaps found in the f.c.c. PMMA crystal. If the particles remain fully connected, an 'inverse opal structure', essentially a negative replica of the PMMA crystal, is obtained when the template is calcinated. If, on the other hand, these connections are lost, the particles become randomly oriented. The authors' result represents a 'third way' between these two extremes.

This happy medium is made possible because the larger and energetically favoured cuboid structures grow, through a process known as Ostwald ripening, at the expense of the smaller spheres, which eventually vanish. Adjacent sheets of the f.c.c. structure are then assumed to collapse on each other, leading to the observed simple-cubic packing. Simple cubic might in fact be the closest — and thus energetically favoured — packing possible for the new arrangement: whereas for spheres, the space-filling factor of 74% for f.c.c. or hexagonal packing is the best that can be achieved, ideal cubes pack closest in tetragonal layers with a space filling of 100%. As Li and colleagues' cuboids are somewhere between a perfect cube and a sphere, it is difficult to predict theoretically what their closest packing would be.

Meanwhile, the interesting question is how general the pathway described by the authors is and what applications the new materials might have. The authors have observed similar ordered regions in a zirconium dioxide–phosphorus pentoxide (ZrO_2–P_2O_5) system, and speculate that the presence of a phosphate-rich phase is crucial to the formation of these regions. Such a phase has a low melting point, and could provide the liquid component necessary for the growth of the cuboid particles at the expense of the rounded ones. The principle involved is quite general, and should be applicable to other systems, with appropriate changes to concentrations, temperatures and times of treatment.

Whether a wider range of structures can be obtained is more difficult to predict. Clearly, a parent structure with differently sized voids is needed. For the formation of a simple-cubic packing, the fact that holes are on top of particles, and vice versa, in adjacent layers of the parent f.c.c. structure, allowing the replica layers to collapse into each other on calcination, seems to be important. Cubes from a redispersed simple-cubic colloidal crystal do not reassemble into an ordered structure.

©2007 Nature Publishing Group

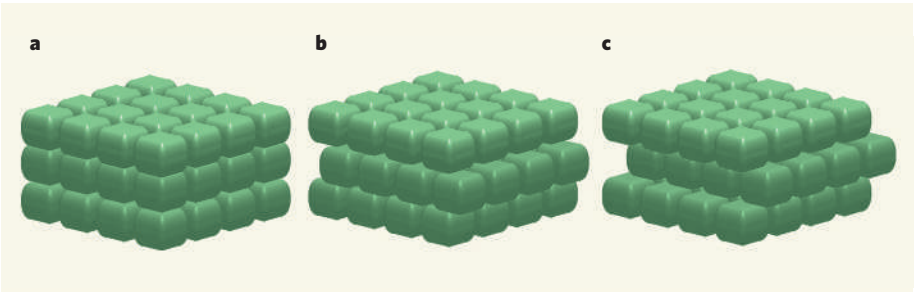


Figure 2 | Packed in stacks. If, as in the case of Li and colleagues' crystals[1], the individual cubes that make up a lattice are not perfect, then packings with the cubes exactly on top of each other (**a**), shifted by half a unit cell along one axis (**b**), and by half a unit cell along two axes (**c**), will all have slightly different packing densities. This will also alter the overall symmetry of the structure, which in turn affects the photonic properties.

Preordering thus has a crucial role. With cuboids deviating from the perfect cube shape, as in Li and colleagues' work, there are structures with higher packing density than the simple cubic packing. An example is a stacking with every second layer shifted, resulting in a 'body-centred' tetragonal (stretched-cube) structure (Fig. 2).

Cuboids that have different surface curvatures could also induce variations in the surface properties and effective cohesion forces of the structure, allowing position-sensitive interactions between nanoparticles that show promise for the design of other structures. If, for instance, Li and colleagues' process[1] could be driven to result in tetrahedral building-blocks, these could form other interesting structures, including some with fewer symmetry elements. This symmetry lowering enables remarkable changes to occur in the photonic bandgap properties that are so essential to the use of these materials as photonic crystals for guiding the propagation of light[7]. The potential of colloidal crystals for photonic applications increases markedly if their features are around the same size as the wavelength of visible light.

The standard approach to generating photonic crystals, rather like the standard techniques for creating patterned circuits in the semiconductor industry, uses lithography, albeit in this case for the generation of three-dimensional structures. Compared with that approach, nanoparticle self-assembly as previously implemented has two great drawbacks: it produces too many defects, and it is essentially restricted to f.c.c. lattices. Li and colleagues' method shows a way to avoid the second drawback. If nanoparticle self-assembly is to have photonic applications, attention must now turn to ridding the simple-cubic packing of old defects inherited from the parent template, and of new ones introduced during the replication process. ■

F. Schüth and F. Marlow are at the Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim, Germany.
e-mail: schueth@mpi-muelheim.mpg.de

1. Li, F., Delo, S. A. & Stein, A. *Angew. Chem. Int. Edn* **46,** 6666–6669 (2007).
2. Marlow, F. & Dong, W. *ChemPhysChem* **4,** 549–554 (2003).
3. Shevchenko, E. V., Talapin, D. V., Kotov, N. A., O'Brien, S. & Murray, C. B. *Nature* **439,** 55–59 (2006).
4. Shevchenko, E. V., Talapin, D. V., Murray, C. B. & O'Brien, S. *J. Am. Chem. Soc.* **128,** 3620–3637 (2006).
5. Leunissen, M. E. *et al. Nature* **437,** 235–240 (2005).
6. Hoogenboom, J. P. *et al. Nano Lett.* **4,** 205–208 (2004).
7. Toader, O. & John, S. *Science* **292,** 1133–1135 (2001).

AUTOIMMUNE DISEASE

Skin deep but complex

Nicole Baumgarth and Charles L. Bevins

The antimicrobial peptide LL37 is essential for normal immune responses to infection or tissue injury. But in the autoimmune skin disorder psoriasis, LL37 propagates disease by forming complexes with host DNA.

The innate (nonspecific) immune system is responsible for detecting pathogens and inducing effector molecules to coordinate subsequent immune responses and combat infection. Genes encoding the receptors of innate immunity are hard-wired in the host genome, and honed to near perfection by the selective pressures of evolution to respond only to non-self targets. When tolerance (non-responsiveness to self) is broken, this can result in autoimmune disease. On page 564 of this issue, Lande *et al.*[1] report a costly glitch in the innate immune response that seems to underlie the development of the common autoimmune skin disease psoriasis*.

Psoriasis is marked by chronic inflammation and excessive proliferation and turnover of the keratinocyte skin cells, which result in characteristic silvery white scaly patches overlying the inflamed skin[2]. The specific causes of psoriasis are unknown, but environmental triggers — including bacterial skin infections, mild trauma and stress — and strong genetic factors underlie the development of the disease. At a cellular level, an accumulation of inflammatory cells largely consisting of activated T cells and antigen-presenting cells, particularly plasmacytoid dendritic cells (pDCs), precedes other aspects of the disease's pathology. Therefore, psoriasis is currently thought of as an autoimmune inflammatory disorder, but what initiates and perpetuates it has remained enigmatic.

Typically, psoriatic lesions contain high levels of several peptide immune mediators such as cytokines, chemokines and antimicrobial peptides. It is likely that these molecules cause the activation and influx of inflammatory cells to, and perhaps other abnormalities in, the skin lesions. For example, previous studies[3,4] had identified high levels of antimicrobial peptides, including β-defensin 2 and cathelicidin LL37, in these lesions.

Cathelicidins are a large family of antimicrobial peptides found in mammals and other vertebrates, and are expressed in both neutrophils and epithelial cells of many tissues[5]. Studies in mice have established an essential role for these peptides in innate immune responses to bacterial infection. Besides protecting the host through their antimicrobial activity, cathelicidins coordinate many actions, including migration of white blood cells and wound-healing.

Humans have only one known cathelicidin, LL37, which contributes to efficient antibacterial defence in psoriasis[4]. But Lande *et al.*[1] describe an alternative activity of this antimicrobial peptide in psoriasis. Normally, the immune mediator interferon-κ (IFN-κ) is the predominant member of the large family of type-I interferons in the skin. In psoriasis, however, its levels are reduced[6], and instead, large amounts of IFN-α are produced locally by pDCs[7], which accumulate in the skin lesions. Increased production of IFN-α is not only central to the development of psoriasis[8] — other autoimmune diseases such as systemic lupus erythematosus have also been linked to the abnormal production and function(s) of this potent immunomodulatory cytokine[9].

Lande *et al.* analysed psoriatic skin extracts for a factor that could elicit IFN-α production by pDCs. They found that LL37 can bind to self DNA, and that the resulting complexes

*This article and the paper concerned[1] were published online on 16 September 2007.

©2007 Nature Publishing Group

NEWS & VIEWS

AUTOIMMUNE DISEASE

Skin deep but complex

Nicole Baumgarth and Charles L. Bevins

The antimicrobial peptide LL37 is essential for normal immune responses to infection or tissue injury. But in the autoimmune skin disorder psoriasis, LL37 propagates disease by forming complexes with host DNA.

The innate (nonspecific) immune system is responsible for detecting pathogens and inducing effector molecules to coordinate subsequent immune responses and combat infection. Genes encoding the receptors of innate immunity are hard-wired in the host genome, and honed to near perfection by the selective pressures of evolution to respond only to non-self targets. When tolerance (non-responsiveness to self) is broken, this can result in autoimmune disease. In a paper published on *Nature*'s website today, Lande *et al.*[1] report a costly glitch in the innate immune response that seems to underlie the development of the common autoimmune skin disease psoriasis.

Psoriasis is marked by chronic inflammation and excessive proliferation and turnover of the keratinocyte skin cells, which result in characteristic silvery white scaly patches overlying the inflamed skin[2]. The specific causes of psoriasis are unknown, but environmental triggers — including bacterial skin infections, mild trauma and stress — and strong genetic factors underlie the development of the disease. At a cellular level, an accumulation of inflammatory cells largely consisting of activated T cells and antigen-presenting cells, particularly plasmacytoid dendritic cells (pDCs), precedes other aspects of the disease's pathology. Therefore, psoriasis is currently thought of as an autoimmune inflammatory disorder, but what initiates and perpetuates it has remained enigmatic.

Typically, psoriatic lesions contain high levels of several peptide immune mediators such as cytokines, chemokines and antimicrobial peptides. It is likely that these molecules cause the activation and influx of inflammatory cells to, and perhaps other abnormalities in, the skin lesions. For example, previous studies[3,4] had identified high levels of antimicrobial peptides, including β-defensin 2 and cathelicidin LL37, in these lesions.

Cathelicidins are a large family of antimicrobial peptides found in mammals and other vertebrates, and are expressed in both neutrophils and epithelial cells of many tissues[5]. Studies in mice have established an essential role for these peptides in innate immune responses to bacterial infection. Besides protecting the host through their antimicrobial activity, cathelicidins coordinate many actions, including migration of white blood cells and wound-healing.

Humans have only one known cathelicidin, LL37, which contributes to efficient antibacterial defence in psoriasis[4]. But Lande *et al.*[1] describe an alternative activity of this antimicrobial peptide in psoriasis. Normally, the immune mediator interferon-κ (IFN-κ) is the predominant member of the large family of type-I interferons in the skin. In psoriasis, however, its levels are reduced[6], and instead, large amounts of IFN-α are produced locally by pDCs[7], which accumulate in the skin lesions. Increased production of IFN-α is not only central to the development of psoriasis[8] — other autoimmune diseases such as systemic lupus erythematosus have also been linked to the abnormal production and function(s) of this potent immunomodulatory cytokine[9].

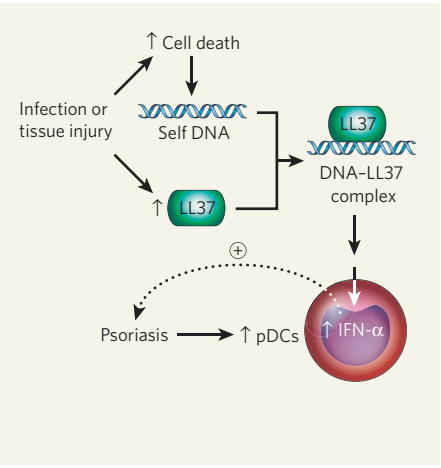


Figure 1 | Role of LL37 in psoriasis. In response to microbial infection or tissue injury, LL37 is produced locally in the skin, where its normal functions include antimicrobial activity and aiding in wound-healing. Lande *et al.*[1] show that LL37 also forms complexes with self DNA that is released from dying cells. In normal skin, these DNA–LL37 complexes probably remain undetected and inconsequential. But in the presence of plasmacytoid dendritic cells (pDCs), which accumulate in the skin lesions of patients with psoriasis, these complexes trigger strong interferon-α (IFN-α) production, which is known to perpetuate the disease[10].

Lande *et al.* analysed psoriatic skin extracts for a factor that could elicit IFN-α production by pDCs. They found that LL37 can bind to self DNA, and that the resulting complexes break tolerance and signal the production of IFN-α (Fig. 1). These findings considerably advance our understanding of psoriasis, and could provide insight into other biological and pathological processes.

But how do the three elements of this response — self DNA, LL37 and pDCs — converge? A high turnover of keratinocytes in the psoriatic lesions, and the associated release of DNA from dying cells, means that high concentrations of human DNA are on hand at the disease site. The high local levels of LL37 can be explained by the pathways that induce its production — infection and tissue injury, which are also known to precede psoriasis onset. Yet the reason for pDC accumulation, a hallmark of psoriasis, is unknown. Could it be that β-defensin 2, the other antimicrobial peptide highly expressed in this disease, attracts pDCs to or retains them in psoriatic lesions? This peptide is known[10] to attract cells that express the chemokine receptor CCR6, and so may recruit CCR6-bearing pDCs[11]. However, as both LL37 and β-defensin 2 are increased in normal skin responses, additional factors specifically associated with psoriasis might lead to an especially robust induction of β-defensin 2 or an exaggerated pDC response.

The deleterious activity of LL37 in psoriatic lesions, as observed by Lande and colleagues, unveils an apparent paradox. It has been shown[12] that microbial stimulation of macrophage cells causes increased expression of the genes encoding the vitamin-D receptor and the enzyme vitamin-D1 hydroxylase. This, in turn, leads to vitamin-D-mediated induction of LL37, which helps macrophages to kill pathogens. If the molecular signalling pathways in skin keratinocytes and macrophages were similar, one might imagine that application of vitamin D to psoriatic lesions would increase LL37 concentrations and perhaps exacerbate the disease. However, the opposite is true. Vitamin-D3 analogues are mainstay drugs for psoriasis[2].

© 2007 Nature Publishing Group

NATURE|AOP|doi:10.1038/nature05893|Published online 16 September 2007

Possible explanations for this puzzle may include tissue-specific effects of vitamin D, altered responses to vitamin D depending on the cytokine milieu in the lesions, and the already maximal induction of LL37 in skin lesions before therapy. One would certainly expect that several biological effects of vitamin D, beyond increasing LL37 expression, may trump its possible effects on LL37. Investigation of this apparent paradox may lead to even more effective therapies for psoriasis.

Abnormal levels of LL37 in the skin have also been linked to other human diseases, including a skin disorder of unknown cause called rosacea[13], and the allergic skin disorder atopic dermatitis, which is associated with susceptibility to skin infections[4,14]. These findings, together with those of Lande *et al.*[1], highlight the importance of tightly controlled LL37 expression for healthy skin. Moreover, that LL37 can bind to self DNA and subsequently activate pDCs may mark a seminal discovery in the field of autoimmunity in general. Earlier work[15] had shown that recognition of complexes of self DNA and antibodies underlies autoimmunity. Thus, complexes of self DNA with components of the immune system, both innate (LL37) and adaptive (antibody), are involved in the perpetuation of autoimmune diseases. Once in complex, self-DNA molecules become strongly immunostimulatory by engaging crucial molecular signalling pathways mediated by Toll-like receptors. The physiological role, if any, of such immunomodulatory complexes remains to be discovered. ■

Nicole Baumgarth is in the Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, The Center for Comparative Medicine and Graduate Group in Immunology, University of California, Davis, California 95616, USA. Charles L. Bevins is in the Department of Microbiology and Immunology, School of Medicine, and Graduate Group in Immunology, University of California, Davis, California 95616, USA.
e-mails: nbaumgarth@ucdavis.edu; clbevins@ucdavis.edu

1. Lande, R. *et al. Nature* doi:10.1038/nature06116 (2007).
2. Schön, M. P. & Boehncke, W.-H. *N. Engl. J. Med.* **352,** 1899–1912 (2005).
3. Harder, J. & Schröder, J.-M. *J. Leukoc. Biol.* **77,** 476–486 (2005).
4. Ong, P. Y. *et al. N. Engl. J. Med.* **347,** 1151–1160 (2002).
5. Zanetti, M. *Curr. Issues Mol. Biol.* **7,** 179–196 (2005).
6. Scarponi, C. *et al. J. Interferon Cytokine Res.* **26,** 133–140 (2006).
7. Liu, Y.-J. *Annu. Rev. Immunol.* **23,** 275–306 (2005).
8. Nestle, F. O. *et al. J. Exp. Med.* **202,** 135–143 (2005).
9. Pascual, V., Banchereau, J. & Palucka, A. K. *Curr. Opin. Rheumatol.* **15,** 548–556 (2003).
10. Yang, D. *et al. Science* **286,** 525–528 (1999).
11. Bendriss-Vermare, N. *et al. Leukemia* **18,** 1491–1498 (2004).
12. Liu, P. *et al. Science* **311,** 1770–1773 (2006).
13. Yamasaki, K. *et al. Nature Med.* **13,** 975–980 (2007).
14. Howell, M. D. *et al. Immunity* **24,** 341–348 (2006).
15. Leadbetter, E. A. *et al. Nature* **416,** 603–607 (2002).

© 2007 Nature Publishing Group

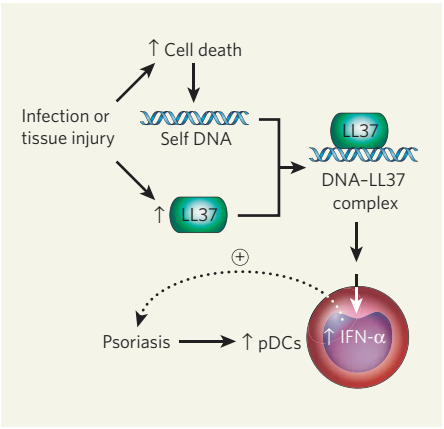


Figure 1 | Role of LL37 in psoriasis. In response to microbial infection or tissue injury, LL37 is produced locally in the skin, where its normal functions include antimicrobial activity and aiding in wound-healing. Lande *et al.*[1] show that LL37 also forms complexes with self DNA that is released from dying cells. In normal skin, these DNA–LL37 complexes probably remain undetected and inconsequential. But in the presence of plasmacytoid dendritic cells (pDCs), which accumulate in the skin lesions of patients with psoriasis, these complexes trigger strong interferon-α (IFN-α) production, which is known to perpetuate the disease[10].

break tolerance and signal the production of IFN-α (Fig. 1). These findings considerably advance our understanding of psoriasis, and could provide insight into other biological and pathological processes.

But how do the three elements of this response — self DNA, LL37 and pDCs — converge? A high turnover of keratinocytes in the psoriatic lesions, and the associated release of DNA from dying cells, means that high concentrations of human DNA are on hand at the disease site. The high local levels of LL37 can be explained by the pathways that induce its production — infection and tissue injury, which are also known to precede psoriasis onset. Yet the reason for pDC accumulation, a hallmark of psoriasis, is unknown. Could it be that β-defensin 2, the other antimicrobial peptide highly expressed in this disease, attracts pDCs to or retains them in psoriatic lesions? This peptide is known[10] to attract cells that express the chemokine receptor CCR6, and so may recruit CCR6-bearing pDCs[11]. However, as both LL37 and β-defensin 2 are increased in normal skin responses, additional factors specifically associated with psoriasis might lead to an especially robust induction of β-defensin 2 or an exaggerated pDC response.

The deleterious activity of LL37 in psoriatic lesions, as observed by Lande and colleagues, unveils an apparent paradox. It has been shown[12] that microbial stimulation of macrophage cells causes increased expression of the genes encoding the vitamin-D receptor and the enzyme vitamin-D1 hydroxylase. This, in turn, leads to vitamin-D-mediated induction of LL37, which helps macrophages to kill pathogens. If the molecular signalling pathways in skin keratinocytes and macrophages were similar, one might imagine that application of vitamin D to psoriatic lesions would increase LL37 concentrations and perhaps exacerbate the disease. However, the opposite is true. Vitamin-D3 analogues are mainstay drugs for psoriasis[2].

Possible explanations for this puzzle may include tissue-specific effects of vitamin D, altered responses to vitamin D depending on the cytokine milieu in the lesions, and the already maximal induction of LL37 in skin lesions before therapy. One would certainly expect that several biological effects of vitamin D, beyond increasing LL37 expression, may trump its possible effects on LL37. Investigation of this apparent paradox may lead to even more effective therapies for psoriasis.

Abnormal levels of LL37 in the skin have also been linked to other human diseases, including a skin disorder of unknown cause called rosacea[13], and the allergic skin disorder atopic dermatitis, which is associated with susceptibility to skin infections[4,14]. These findings, together with those of Lande *et al.*[1], highlight the importance of tightly controlled LL37 expression for healthy skin. Moreover, that LL37 can bind to self DNA and subsequently activate pDCs may mark a seminal discovery in the field of autoimmunity in general. Earlier work[15] had shown that recognition of complexes of self DNA and antibodies underlies autoimmunity. Thus, complexes of self DNA with components of the immune system, both innate (LL37) and adaptive (antibody), are involved in the perpetuation of autoimmune diseases. Once in complex, self-DNA molecules become strongly immunostimulatory by engaging crucial molecular signalling pathways mediated by Toll-like receptors. The physiological role, if any, of such immunomodulatory complexes remains to be discovered. ■

Nicole Baumgarth is in the Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, The Center for Comparative Medicine and Graduate Group in Immunology, University of California, Davis, California 95616, USA. Charles L. Bevins is in the Department of Microbiology and Immunology, School of Medicine, and Graduate Group in Immunology, University of California, Davis, California 95616, USA.
e-mails: nbaumgarth@ucdavis.edu; clbevins@ucdavis.edu

1. Lande, R. *et al. Nature* **449,** 564–569 (2007).
2. Schön, M. P. & Boehncke, W.-H. *N. Engl. J. Med.* **352,** 1899–1912 (2005).
3. Harder, J. & Schröder, J.-M. *J. Leukoc. Biol.* **77,** 476–486 (2005).
4. Ong, P. Y. *et al. N. Engl. J. Med.* **347,** 1151–1160 (2002).
5. Zanetti, M. *Curr. Issues Mol. Biol.* **7,** 179–196 (2005).
6. Scarponi, C. *et al. J. Interferon Cytokine Res.* **26,** 133–140 (2006).
7. Liu, Y.-J. *Annu. Rev. Immunol.* **23,** 275–306 (2005).
8. Nestle, F. O. *et al. J. Exp. Med.* **202,** 135–143 (2005).
9. Pascual, V., Banchereau, J. & Palucka, A. K. *Curr. Opin. Rheumatol.* **15,** 548–556 (2003).
10. Yang, D. *et al. Science* **286,** 525–528 (1999).
11. Bendriss-Vermare, N. *et al. Leukemia* **18,** 1491–1498 (2004).
12. Liu, P. *et al. Science* **311,** 1770–1773 (2006).
13. Yamasaki, K. *et al. Nature Med.* **13,** 975–980 (2007).
14. Howell, M. D. *et al. Immunity* **24,** 341–348 (2006).
15. Leadbetter, E. A. *et al. Nature* **416,** 603–607 (2002).

IMAGING TECHNOLOGY

Harmonic pictures in a flash

John Spence

Making films of atomic-scale processes as they happen makes huge demands on any imaging system. One approach combines the advantages of pulsed laser harmonics and computerized image reconstruction.

To make a movie of molecular rearrangements, the criteria are strict. Enough of the illuminating beam should be scattered off the imaged object during each frame to form an image (at least one particle for each pixel), so the beam must be very intense and have a high scattering probability. This scattering should also be mostly elastic, and not transfer too much damaging energy to the sample. But most crucially, the imaging camera should have a frame speed of a few femtoseconds, coupled with atomic-scale spatial resolution.

These dual demands rule out most conventional imaging techniques. Optical lasers, for example, can offer the right sort of speed (a single period of laser light lasts about 2.5 fs), but they fall down on spatial resolution. (Spatial resolution is generally limited to about the wavelength of the probe, and the wavelength of optical light lies in the region of 400–700 nanometres.) Conversely, electrons have a small enough wavelength, but lack the requisite speed owing to the added complication of their charge interactions. Meanwhile, X-rays are limited by the aberrations and fabrication difficulties of the 'zone-plate' lenses that focus them.

Writing in *Physical Review Letters*, Sandberg *et al.*[1] report taking a valuable step towards circumventing these problems. They combine recent breakthroughs in lensless imaging — which avoids the problems associated with lens aberrations by using a computer for image reconstruction — with advances in laser-driven X-ray generation to overcome the problem of

©2007 Nature Publishing Group

NATURE|Vol 449|4 October 2007

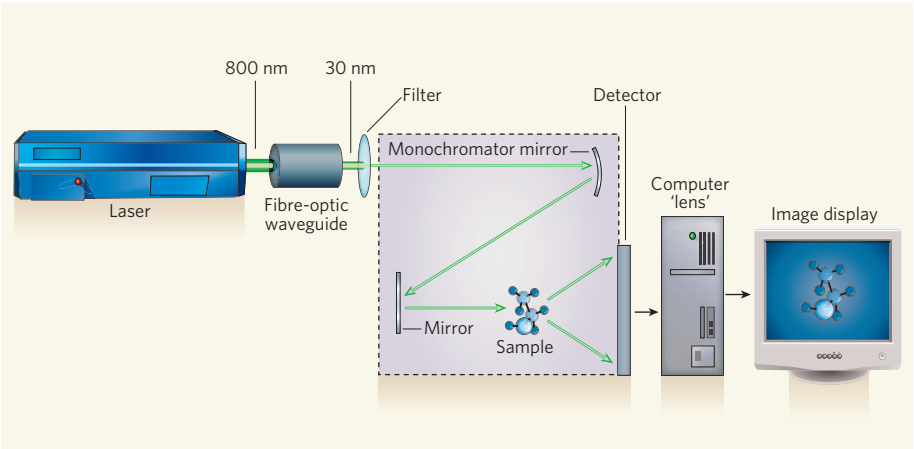


Figure 1 | Fast pictures without lenses. Sandberg *et al.*[1] convert laser light at an optical wavelength into 30-nm-wavelength extreme ultraviolet pulses in an argon-filled fibre-optic waveguide with an inner diameter of 150 μm. The resulting fast, directed 'high-harmonic' beam has a diameter of 25 μm that is further reduced in intensity (by several orders of magnitude) by filters and a monochromator. The authors get around the need for low-wavelength lenses (which are prone to large aberrations) by using an iterative phase-retrieval computer algorithm to extract the maximum information from the scattered light.

spatial resolution, while preserving the laser's innate speed.

This advance is just the latest act in a fascinating story of the replacement of lenses by computers in imaging technology. The origins lie in the realms of signal processing, X-ray crystallography and electron microscopy[2], and the breakthrough for X-rays came in 1999, with the first non-holographic reconstruction by numerical means of an image made by scattering X-rays from a non-periodic sample[3]. The current state-of-the-art[4] fast, lensless imaging technique uses radiation produced by a free-electron laser at a synchrotron facility to make, in a single shot, images with a temporal resolution of 25 fs and a spatial resolution of 90 nm.

The secret behind all these techniques is an iterative phase-retrieval algorithm[2]. Iterative phase retrieval is one answer (various forms of holography and X-ray crystallography use other approaches) to the notorious 'phase problem' — that all detectors record only the intensity of the radiation that impinges on them, throwing away the phase information. Under suitable experimental conditions, however, this phase information is encoded in the intensity, and may be recovered if the intensity is sampled correctly. The algorithms iterate between the image and the scattering pattern (which are related by a mathematical operation known as a Fourier transform), while imposing known information, such as the approximate boundary of the object, on each. The great strength of such an algorithm is that it can be implemented for any type of imaging particle of any wavelength. Each particle interacts differently with a sample, and so can potentially provide new information about it. On the downside, such algorithms introduce constraints on the sample geometry, and coherence and aberrations in the illuminating wavefield become important.

Sandberg *et al.*[1] generate very 'soft' X-rays (actually, extreme ultraviolet radiation) with a wavelength of about 30 nm by scattering intense pulses of infrared laser light of a much longer wavelength (800 nm) on gas atoms (Fig. 1). These X-rays scatter from the object, and are combined into an image with 214-nm resolution using a phase-retrieval algorithm. The imaging technique exploits high-frequency harmonics produced when laser light of energy hw (w is the laser frequency and h is Planck's constant) passes through a nonlinear medium. An atomic electron in the medium absorbs n laser photons before spitting out a single high-energy photon of n times the energy (and a similarly increased frequency nw), but the same properties of phase coherence and pulse duration as the driving laser.

Classically, we can think of the atomic electron being initially ejected by the laser pulse, before being returned to the atom during the second half of the laser cycle when the electric field reverses direction. The resulting acceleration produces radiation (bremsstrahlung) at the high-harmonic frequency. Importantly, this beam of radiation is directed forwards, and its phase coherence and conversion efficiency are greatly enhanced if generated inside a waveguide. High harmonics extending into soft X-ray frequencies were first observed[5] in 1988, and used 5 years ago for holographic imaging with a resolution of about 10 μm (ref. 6).

Might a high-harmonic technique such as that of Sandberg's group one day provide competition for the large synchrotron particle accelerators currently used for molecular crystallography and the like? Synchrotrons provide tunable radiation with wavelengths from tens of nanometres to less than a tenth of a nanometre by collecting the bremsstrahlung from high-energy electrons accelerated over an optimized path. The wavelengths at which the high-harmonic technique is viable are continuing to fall (a collisional X-ray laser of wavelength 13 nm seeded by high-harmonic radiation is on the cards), pulsing rates are increasing and pulse duration is decreasing. A similar scheme using laser standing waves to undulate electron beams produces tunable, directed 35-kiloelectronvolt bremsstrahlung X-rays (which have wavelengths of a few hundredths of a nanometre). This would be useful for protein crystallography, but the apparatus occupies a room rather than a table. Then there is wakefield acceleration, in which laser pulses running through a plasma are used to accelerate electrons to gigaelectronvolt energies over a few centimetres.

This cornucopia of techniques is starting to produce viable competitors of synchrotrons in the effort to obtain higher-resolution, faster images. But the competition between the alternative techniques is intense. Sandberg and colleagues' method[1], although promising, has some way to go. Its spatial resolution is still not sharp enough to see atoms, and the images required one and a half hours' exposure time with continuous 25-fs pulsing, owing to losses in the optics. One way around this problem for inorganic samples, in which repetitive processes such as electronic excitation and atomic motion can be triggered by another synchronized laser, is 'stroboscopic' imaging, which builds up pictures at different instants during the repeated cycle. In this case, fewer scattered particles are needed in each pulse, because many noisy images can be added together. A moving picture can thus be constructed by varying the delay between images. With biological samples, the limiting factor for this technique is the radiation damage caused by high exposure.

Despite these outstanding problems, lensless imaging in biology and materials science using electrons, neutrons and X-rays spans a wide and increasing array of techniques and capabilities. It is still early days — atoms were first seen[7] in the field-ion microscope in 1951, and soon after with electron microscopy, whereas atomic-resolution lensless images of a single carbon nanotube were first reconstructed[8] in 2003. One viable proposal for a non-damaging, sub-femtosecond, atomic-scale imaging technique based on self-diffraction of high-harmonic electrons from laser-aligned gas molecules already exists[9]. That is indeed a goal worth striving for. ■

John Spence is in the Department of Physics, Arizona State University, Tempe, Arizona 85287-1504, USA.
e-mail: spence@asu.edu

1. Sandberg, R. L. *et al. Phys. Rev. Lett.* **99,** 098103 (2007).
2. Spence, J. C. H. in *Science of Microscopy* (eds Hawkes, P. W. & Spence, J. C. H.) 1196-1227 (Springer, New York, 2007).
3. Miao, J., Charalambous, P., Kirz, J. & Sayre, J. *Nature* **400,** 342-344 (1999).
4. Chapman, H. *et al. Nature Phys.* **2,** 839-843 (2006).
5. Ferray, M. *et al. J. Phys. B* **21,** L31-L35 (1988).
6. Bartels, R. A. *et al. Science* **297,** 376-378 (2002).
7. Müller, E. W. *Z. Phys.* **31,** 136-139 (1951).
8. Zuo, J. M., Vartanyants, I., Gao, M., Zhang, R. & Nagahara, L. A. *Science* **300,** 1419-1421 (2003).
9. Corkum, P. B. & Krausz, F. *Nature Phys.* **3,** 381-387 (2007).

©2007 Nature Publishing Group

CHEMICAL BIOLOGY

More charges against aggregation

Michele Vendruscolo and Christopher M. Dobson

Protein aggregation causes problems for biotechnology and leads to many fatal human diseases. But a grasp of the physical principles involved enables 'superproteins' to be designed that have exceptional solubilities.

Proteins evolved under the stringent conditions imposed by the cellular environment. This means that, out of the vast number of possible amino-acid sequences in proteins, only a tiny fraction actually occur in nature. The advent of protein engineering[1] provides an opportunity to create proteins that have sequences that have never been found in living organisms, and that could have properties unparalleled in their natural counterparts. Reporting in the *Journal of the American Chemical Society*, Liu and colleagues[2] show that it is possible to design functional proteins with very high electrostatic charge that turn out to be far more soluble in water than their naturally occurring analogues.

Although a detailed knowledge of the biological factors that influence the behaviour of proteins is crucial to an in-depth understanding of their fundamental nature, considerable insights can also be gained by examining their physical and chemical properties. This idea has been reinforced by the discovery[3] that such properties are closely linked to the tendency of proteins to aggregate into non-functional polymeric structures, often known as amyloid assemblies. These structures are best known as the basis for the plaques that form in Alzheimer's disease, and they, or their precursors, can be highly toxic[3]. The need to avoid aggregation has limited amino-acid sequences to those that yield proteins with a relatively narrow range of specific physical attributes — such as hydrophobicity and net electrostatic charge.

Yet natural proteins have overcome many of these limitations to exhibit a wide range of solubilities. Such versatility enables them to work in diverse environments — from non-polar lipid membranes to the acid-bath of the stomach — and suggests that altering their solubility even more radically could result in artificial, 'high-performance' proteins. Moreover, if enzymes can be made to function in inorganic solvents other than water[4], and in organic solvents other than the lipids that make up biological membranes[5], new catalytic roles might emerge. This type of research has great potential for the chemical industry, and already allows carefully engineered proteins to be self-assembled into technologically valuable materials in the laboratory[6].

It follows, too, that a promising therapeutic strategy[7] to combat protein-deposition disorders such as Alzheimer's disease is to produce slightly more soluble versions of the proteins whose aggregation is the root cause of the problem. Such modified proteins will reduce the tendency of their natural counterparts to aggregate, while remaining compatible with their cellular environment[8]. This strategy might be aided by the actions of molecular 'chaperones' that protect the mutated proteins and promote their safe interactions with their environment[9].

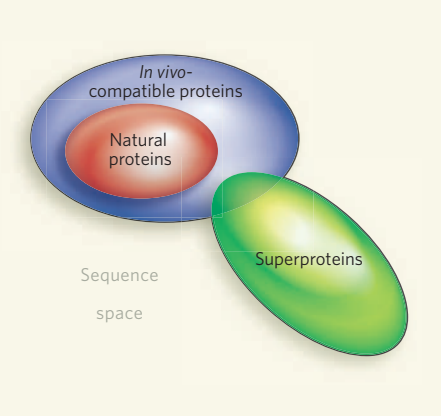


Figure 1 | An exploration chart of protein amino-acid sequences. The number of possible amino-acid sequences that can make proteins of the size found in living systems is enormous (larger than 10^{400}), and can be represented as the 'sequence space'. The number of naturally occurring proteins is much smaller (probably fewer than 10^{12}); such proteins have broadly similar physical properties, and form a cluster (red oval) in sequence-space. They are surrounded (purple region) by sequences that have not been found in nature, but whose properties might not differ much from those of natural proteins. 'Superproteins' are now being engineered that ignore at least some biologically imposed restrictions on amino-acid sequence, and these occupy regions of sequence space (green oval) unexplored by natural proteins. Such molecules can be designed to have unusual properties, such as Liu and colleagues' highly charged proteins[2], which resist aggregation.

With all of this in mind, Liu and colleagues[2] set out to increase the solubility of proteins dramatically. They began by modifying green fluorescent protein (GFP), which is widely used as an optical reporter for monitoring cellular processes. Using protein engineering, the authors produced GFP variants with a net charge ranging from –30 to +46; for comparison, the net charge of most natural proteins is in the range of –10 to +10. By clever design, these supercharged versions of GFP not only maintained their structural stability *in vitro*, but also remained soluble when exposed to conditions that normally cause proteins to aggregate — such as heating to high temperatures, or treatment with a chemical additive that causes the protein to denature. Liu and colleagues[2] then went on to modify other proteins, and found that the process of supercharging can be achieved without altering the proteins' normal functions.

The authors' results[2] are impressive. It will be fascinating to explore the extent to which such radically altered proteins can avoid unfavourable interactions *in vivo* with other cellular components, interactions that could result in toxicity. Even if such events occur, an armoury of additional design tools is available to modify the proteins further to provide therapeutic compounds.

More generally, the evidence from this study[2] suggests that the ability to use protein engineering and design techniques in new ways to sample largely unexplored amino-acid sequences holds great promise for applications in medicine, biotechnology and even materials science. These opportunities arise from the recognition that it may be possible to overcome the stringent limitations that evolution has imposed on the physical and chemical properties of naturally occurring proteins. ■

Michele Vendruscolo and Christopher M. Dobson are in the Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

1. Fersht, A. R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York, 1999).
2. Lawrence, M. S., Phillips, K. J. & Liu, D. R. *J. Am. Chem. Soc.* **129,** 10110–10112 (2007).
3. Chiti, F. & Dobson, C. M. *Annu. Rev. Biochem.* **75,** 333–366 (2006).
4. van Rantwijk, F. & Sheldon, R. A. *Chem. Rev.* **107,** 2757–2785 (2007).
5. Klibanov, A. M. *Nature* **409,** 241–246 (2001).
6. Zhang, S. *Nature Biotechnol.* **21,** 1171–1178 (2003).
7. Lansbury, P. T. *Nature Biotechnol.* **19,** 112–113 (2001).
8. Tartaglia, G. G., Pechmann, S., Dobson, C. M. & Vendruscolo, M. *Trends Biochem. Sci.* **32,** 204–206 (2007).
9. Queitsch, C., Sangster, T. A. & Lindquist, S. *Nature* **417,** 618–624 (2002).

Correction

In the News & Views article "Atomic physics: A whiff of antimatter soup" by Clifford M. Surko (*Nature* 449, 153–155; 2007), the first prediction of the positronium atom is ascribed to John Wheeler in 1946. In fact the story, as recounted for example by Helge Kragh (H. Kragh *J. Chem. Educ.* 67, 196–197; 1990), is more complex: Stjepan Mohorovičić proposed a similar 'electrum' atom in 1934, although the antiparticles of his system were not the positrons of quantum theory; and Arthur E. Ruark, in work published in 1945 and at the time unknown to Wheeler, coined the term 'positronium' for "an unstable atom composed of a positron and a negative electron".

In the same News & Views article, it is also stated that a γ-ray laser could be made from a Bose–Einstein condensate of positronium molecules (Ps_2); in fact, positronium atoms (Ps) would be the starting point.

©2007 Nature Publishing Group

ARTICLES

Mesenchymal stem cells within tumour stroma promote breast cancer metastasis

Antoine E. Karnoub[1], Ajeeta B. Dash[2], Annie P. Vo[1], Andrew Sullivan[2], Mary W. Brooks[1], George W. Bell[1], Andrea L. Richardson[3], Kornelia Polyak[4], Ross Tubo[2] & Robert A. Weinberg[1]

Mesenchymal stem cells have been recently described to localize to breast carcinomas, where they integrate into the tumour-associated stroma. However, the involvement of mesenchymal stem cells (or their derivatives) in tumour pathophysiology has not been addressed. Here, we demonstrate that bone-marrow-derived human mesenchymal stem cells, when mixed with otherwise weakly metastatic human breast carcinoma cells, cause the cancer cells to increase their metastatic potency greatly when this cell mixture is introduced into a subcutaneous site and allowed to form a tumour xenograft. The breast cancer cells stimulate *de novo* secretion of the chemokine CCL5 (also called RANTES) from mesenchymal stem cells, which then acts in a paracrine fashion on the cancer cells to enhance their motility, invasion and metastasis. This enhanced metastatic ability is reversible and is dependent on CCL5 signalling through the chemokine receptor CCR5. Collectively, these data demonstrate that the tumour microenvironment facilitates metastatic spread by eliciting reversible changes in the phenotype of cancer cells.

The origins of the invasive and metastatic phenotypes of carcinoma cells have been the subjects of intense investigation. Whereas some current models depict these phenotypes as cell-autonomous alterations specified by the genomes of cancer cells, alternative views propose that metastatic traits are acquired through exposure of epithelial cancer cells to paracrine signals that they receive from mesenchymal cell types within the tumour-associated stroma. Although several lines of evidence demonstrate the contributions of stromal cells to primary tumour growth[1], direct experimental demonstration of the influence of these various cells on the metastatic abilities of cancer cells has been difficult to obtain. This is due, in part, to the complexity of the mesenchymal cell types that are recruited into the stroma, and to the elusive nature of the putative paracrine signals that are exchanged between the mesenchymal and epithelial compartments of a tumour. Recent reports proposed that the bone-marrow-derived mesenchymal stem cell (MSC) is a cell type that is recruited in large numbers to the stroma of developing tumours[2]. To characterize better the role of this stromal cell in tumorigenesis, we set out to determine whether MSCs could supply contextual signals that serve to promote cancer metastasis.

Mesenchymal stem cells are pluripotent progenitor cells that contribute to the maintenance and regeneration of a variety of connective tissues, including bone, adipose, cartilage and muscle[3]. Although MSCs reside predominantly in the bone marrow, they are also distributed throughout many other tissues, where they are thought to serve as local sources of dormant stem cells[4,5]. The contributions of MSCs to tissue formation become apparent only in cases of tissue remodelling after injury or chronic inflammation. These conditions are typically accompanied by the release of specific endocrinal signals from the injured or inflamed tissue that are then transmitted to the bone marrow, leading to the mobilization of multi-potent MSCs and their subsequent recruitment to the damage site[6]. For example, MSCs have been shown to contribute to the formation of fibrous scars after injury[7].

The formation of breast carcinomas is often accompanied by a well-orchestrated desmoplastic reaction, which involves the recruitment of a variety of stromal cells with both pro- and anti-tumorigenic activities[1]. Such response closely resembles wound healing and scar formation, and entails the constant deposition of growth factors, cytokines and matrix-remodelling proteins that render the tumour site a 'wound that never heals'[8]. This suggests that, similar to sites of injury, actively growing tumours recruit MSCs through the release of various endocrine and paracrine signals. Indeed, as we have found, mouse stroma prepared from developing human MCF7/Ras or MDA-MB-231 breast cancer xenografts is rich in cells with an ability to generate fibroblastoid colony-forming units (CFU-F) *in vitro* (Supplementary Fig. 1a), a hallmark of MSCs[3]. The absence of such colonies from control Matrigel plugs or from neighbouring tissues (negative control; Supplementary Fig. 1a) suggested that endogenous murine MSCs localize specifically to sites of neoplasia.

To investigate whether human breast cancer cells also have the ability to attract human MSCs, we established a transwell assay in which bone-marrow-derived human MSCs were allowed to migrate towards media derived from MCF7/Ras or MDA-MB-231 cultures. We found that human MSCs migrated much more avidly (~11-fold more) towards media derived from these cancer cells than towards control media (Supplementary Fig. 1b). More importantly, green fluorescent protein (GFP)-labelled human MSCs infused into the venous circulation of mice bearing MCF7/Ras or MDA-MB-231 human breast cancer xenografts localized specifically to the developing tumours, with no observable accumulation in other tissues, such as the kidneys (Supplementary Fig. 1c), liver and spleen (data not shown). Such findings indicated that MSCs are specifically recruited by subcutaneous breast xenografts, and corroborated recent studies that described the localization of systemically infused MSCs to other types of malignancy, such as gliomas[9,10], colon carcinomas[11,12], ovarian carcinomas[13], Kaposi's sarcomas[14] and melanomas[15].

[1]Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [2]Genzyme Corporation, Framingham, Massachusetts 01701, USA. [3]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [4]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.

©2007 Nature Publishing Group

MSCs enhance breast cancer metastasis

To investigate the functional consequences of the heterotypic interactions between MSCs and mammary carcinoma cells, we established a xenograft model in which GFP-labelled MCF7/Ras, MDA-MB-231, MDA-MB-435 and HMLER (see Methods) human breast cancer cells (BCCs) were mixed with bone-marrow-derived human MSCs (hereafter referred to as MSCs) and injected subcutaneously into immunocompromised mice. The growth kinetics of the MSC-containing tumours (BCCs plus MSCs) were compared to those of BCCs injected alone (BCCs) over the subsequent 8–12 weeks, after which the histopathology of the resulting tumours was studied.

We found that MSCs accelerated the growth of MCF7/Ras tumours without affecting the kinetics of MDA-MB-231-, MDA-MB-435- or HMLER-containing tumours (Fig. 1a). More importantly, whereas mice carrying tumours composed only of BCCs exhibited few microscopic metastases in the lungs (Fig. 1b, d), mice bearing the mixed MCF7/Ras+MSC, MDA-MB-231+MSC, MDA-MB-435+MSC and HMLER+MSC tumours displayed a marked increase in the numbers of micro- and macroscopic lung metastases (Fig. 1b, d). Normalized counts of the metastatic nodules in the lungs of BCC+MSC-bearing mice compared to their BCC-control littermates revealed two-, three-, four- and sevenfold enhancements in the overall numbers of detectable HMLER, MDA-MB-435, MCF7/Ras and MDA-MB-231 metastatic deposits, respectively (Fig. 1c). Furthermore, in contrast to the MDA-MB-231-bearing mice, the MDA-MB-231+MSC-bearing mice showed metastases to various other tissues, including the mammary glands (Supplementary Table 1). Although all four of the tested cell lines exhibited enhanced metastatic potential after admixture of MSCs, we chose to focus further analysis on the MDA-MB-231 tumour model, because it displayed the greatest relative increase in MSC-induced metastasis without any concomitant effect on either tumour cell proliferation (as revealed by Ki67 staining; Supplementary Fig. 2) or overall primary tumour growth kinetics.

We note that admixture of other types of mesenchymal cells—specifically WI-38 or BJ human fibroblasts (Supplementary Fig. 3 and data not shown)—to MDA-MB-231 cancer cells before injection into host mice did not result in either enhanced growth kinetics (Supplementary Fig. 3a, b) or increased numbers of lung metastases (Supplementary Fig. 3c, d). Taken together, these observations indicated that the metastasis-enhancing powers were a specific property of admixed MSCs or derivatives thereof.

Reversible metastasis

Implantation of MSCs either contralaterally to MDA-MB-231 cells or in nearby separate sites of injection did not affect the metastatic potential of the resulting primary tumours (data not shown), indicating that MSCs could enhance cancer metastasis only when they were in close proximity to the engrafted BCCs. This influence might be ascribed to various effects that MSCs exert on the commingled carcinoma cells. Thus, the MSCs might favour the outgrowth of rare variants within the MDA-MB-231 cell populations that exhibit unusually high metastatic powers. Alternatively, the MSCs might cause otherwise weakly metastatic MDA-MB-231 cells to acquire enhanced metastatic abilities. This latter mechanism suggests the possibility that the acquisition of the metastatic phenotype might be reversible, in that carcinoma cells might revert to a lower metastatic state once they were no longer in close contact with MSCs.

To resolve between these two mechanisms, explants of MDA-MB-231 cells were prepared from BCC plus MSC primary tumours (T-explants) as well as from their derived lung metastases (L-explants), expanded *in vitro*, cleared from contaminating stromal components, and then re-injected into subcutaneous sites in host mice in order to evaluate their respective metastatic powers (Fig. 2a). Although the growth rate of the resulting L-explant primary tumours was marginally enhanced compared to their T-explant counterparts (Fig. 2b, c), these L-explant cells were no more metastatic than the parental T-explant cancer cells (Fig. 2d). This suggested that the

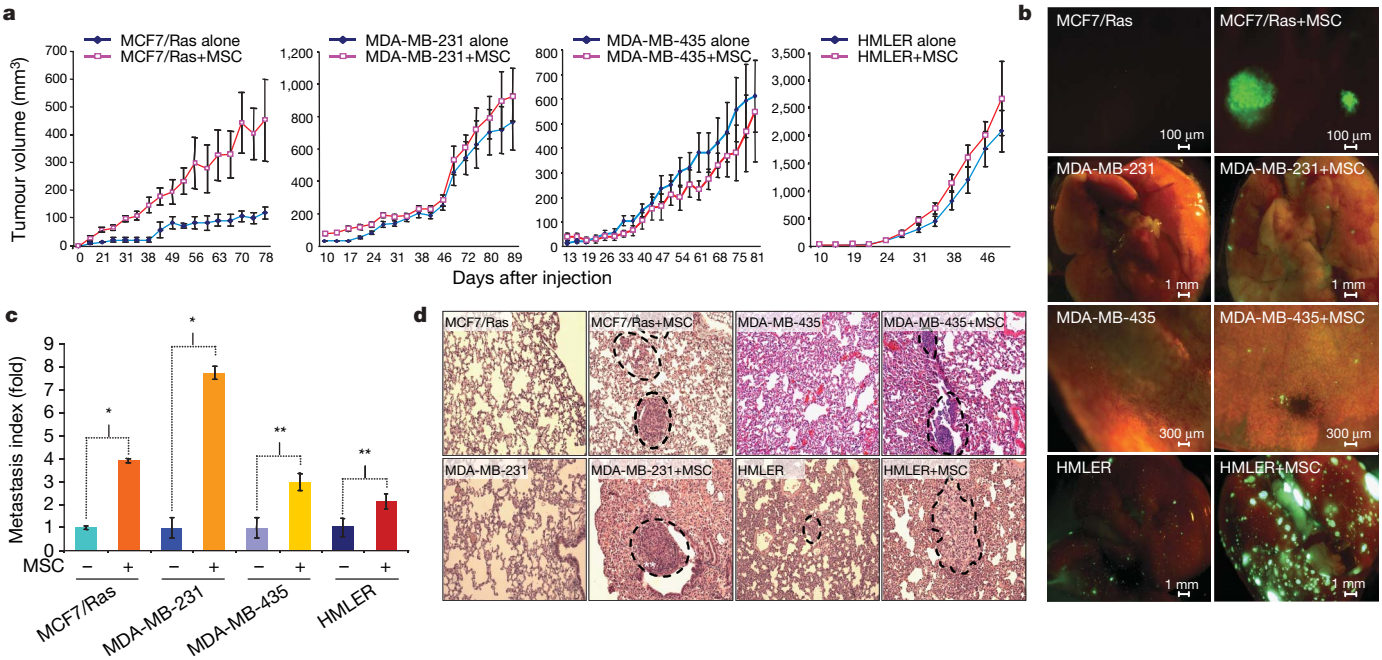


Figure 1 | MSCs promote breast cancer metastasis. **a**, Tumour volume measurements (mean ± s.e.m.) of 500,000 GFP-labelled BCCs injected subcutaneously into nude mice with or without 1.5×10^6 MSCs. Representative data from multiple experiments are shown. Diamonds, BCCs alone, $n = 5$–7 mice per group; squares, BCCs plus MSCs, $n = 5$–8 mice per group. **b**, Representative bright-field/fluorescence images of lungs of mice bearing the indicated tumours. Cancer colonies are in green. MCF7/Ras-bearing mice were killed at approximately day 150 to allow these tumours to grow to comparable sizes to their MCF7/Ras+MSC counterparts. **c**, The lung metastasis indices pooled within each cohort of mice in **a** are expressed as fold increase (±s.e.m.) over controls. Data shown are representative of multiple repeats. Asterisk, $P < 0.01$, double asterisk, $P < 0.05$ using one-tailed Student's *t*-test. **d**, Representative haematoxylin-and-eosin-stained sections of lungs of mice bearing the indicated tumours. Metastases are delineated by a dashed line.

©2007 Nature Publishing Group

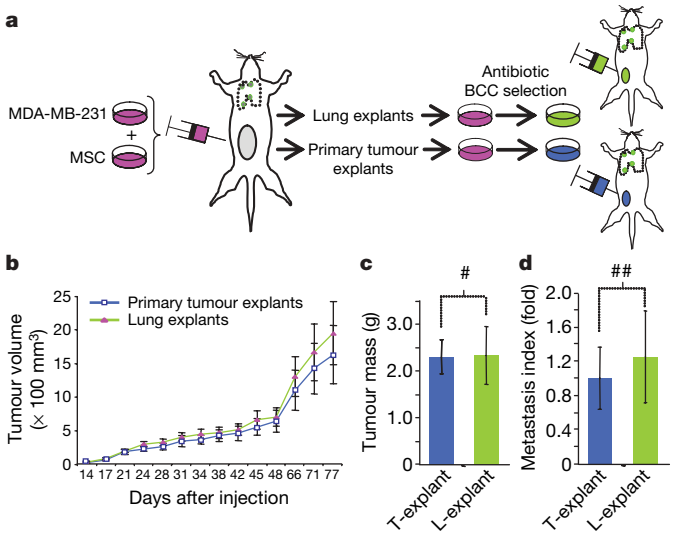


Figure 2 | MSC-induced increase in the metastasis of MDA-MB-231 cells involves reversible mechanisms. **a**, BCCs were recovered from lung or primary tumour tissues, cleared of stromal contaminants by culture in blasticidin-containing media (5 μg ml^{-1}), and re-injected as primary subcutaneous tumours in recipient animals. **b**, Tumour growth (means ± s.e.m.) of 500,000 GFP-labelled lung-derived (L-explant) or primary tumour-derived (T-explant) MDA-MB-231 cells inoculated subcutaneously. Data shown are representative of multiple independent experiments in which four different paired batches of L-explant and T-explant cultures were assayed in parallel. MDA-MB-231-T-explant ($n = 8$ mice); MDA-MB-231-L-explant ($n = 10$ mice). **c**, Masses (means ± s.e.m.) of tumours in **b**. Hash, $P > 0.4$ using one-tailed Student's *t*-test and indicates no statistical significance. **d**, Lung metastasis index of mice in **c**. Double hash, $P > 0.3$ using one-tailed Student's *t*-test and indicates no statistical significance.

MSC-induced metastatic powers reflected a reversibly induced trait of the MDA-MB-231 cells, and that the ability of these cells to metastasize to the lungs was a consequence of their 'education' by MSCs in the primary tumour rather than the selection of rare variants of MDA-MB-231 cells that display elevated metastatic potency in a stable fashion.

The effects that the MSCs exerted on the BCCs might have occurred within the site of primary tumour formation. Alternatively, the MSCs might have accompanied the metastasizing BCCs to sites of metastasis formation. To distinguish between these two possibilities, we admixed ds-red-labelled MSCs to GFP-labelled MDA-MB-231 cells and implanted the mixture subcutaneously in host mice. We found that the tumour-derived lung metastases contained green-labelled MDA-MB-231 cells but no detectable red-labelled MSCs (or their derivatives; Supplementary Fig. 4a) when scored 4, 5 or 6 weeks after primary tumour implantation. The absence of red-labelled MSCs from the lung metastatic sites cannot be ascribed to an inhospitable lung parenchyma, as MSCs that lodge in the lungs of recipient animals after tail-vein infusion survive in that environment for ~6 weeks after injection (Supplementary Fig. 4b). Hence, it appeared that the admixed MSCs do not migrate in large numbers to the sites of metastasis, and that they exerted their pro-metastatic effects on BCCs in the context of primary tumours.

CCL5 in MSC-induced metastasis

The aforementioned observations indicate that MSCs supply locally acting paracrine cues that induce BCCs within primary tumours to metastasize. To understand this crosstalk better, *in vitro* co-cultures of MDA-MB-231 breast cancer cells and MSCs were established and their conditioned media were screened for the levels of various cytokines, chemokines and growth factors using the Luminex-based Bio-Plex suspension array system (Fig. 3a). In some cases, the resulting

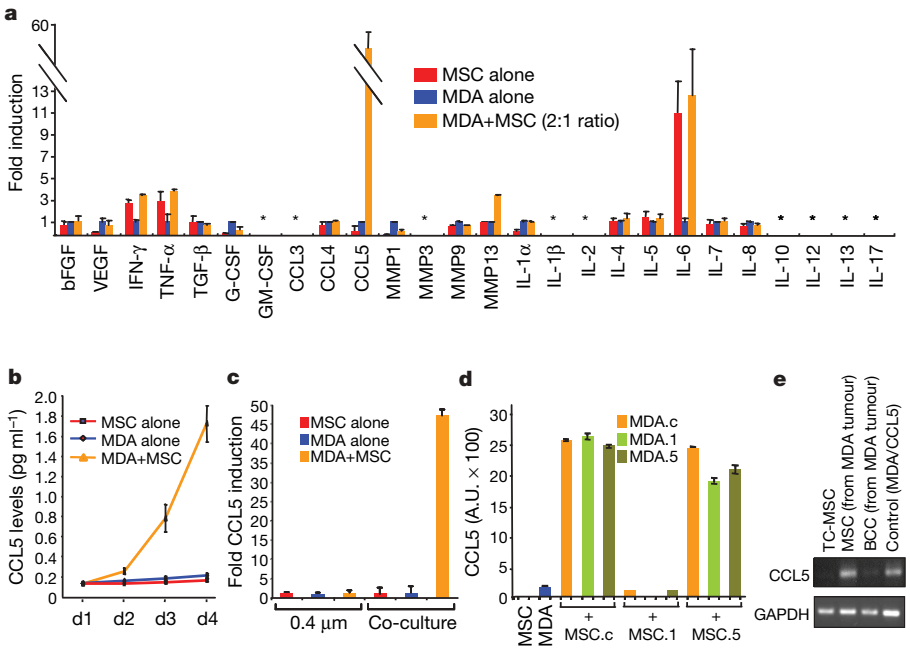


Figure 3 | The interaction of BCCs with MSCs causes a rise in the levels of CCL5. **a**, MDA-MB-231, MSCs, or MDA-MB-231+MSCs were cultured in complete media for 3 days. The levels of various factors in the cell-free culture supernatants were measured by xMAP Bio-Plex cytokine arrays at day 3, and were normalized to the levels observed in the media of BCCs cultured alone. Data are expressed as fold induction ± s.d. of triplicates. Asterisk indicates undetectable levels. **b**, CCL5 ELISA on the media of MDA-MB-231, MSCs, or MDA-MB-231+MSC cultures (1:3 MDA:MSCs) at the indicated time points. Data points represent means ± s.d. of quadruplicates. **c**, BCCs were separated from co-cultured MSCs by a 0.4-μm membrane. CCL5 levels were probed by ELISA on the culture supernatants. Data are expressed as fold induction over levels seen in MDA-MB-231 culture supernatants (mean ± s.d. of triplicates). **d**, CCL5 ELISA on the supernatants of MSC-siluc (MSC.c), MSC-siCCL5.1 (MSC.1) and MSC-siCCL5.5 (MSC.5) co-cultured with MDA-MB-231-siluc (MDA.c), MDA-MB-231-siCCL5.1 (MDA.1), or MDA-MB-231-siCCL5.5 (MDA.5). Data are expressed as means ± s.d. of triplicates in arbitrary units (A.U.). **e**, RT–PCR analyses of CCL5 in MSCs and BCCs sorted from GFP–MSC+MDA-MB-231 tumours (3:1 ratio) 4 weeks after tumour implantation. Tissue-cultured MSCs (TC-MSC) and MDA-MB-231/CCL5 cells were used as controls. GAPDH was used for equal loading.

©2007 Nature Publishing Group

levels of certain released factors (for example, interferon-γ or tumour-necrosis factor-α) reflected the additive contributions of the two cell types when cultured on their own. Notably, the levels of only one cytokine, CCL5, reflected a synergistic interaction between the MSCs and BCCs, as it accumulated to levels ~60-fold higher than those produced by pure BCC cultures (Fig. 3a). This cooperative induction of CCL5 was proportional to the numbers of MSCs mixed with the BCCs (Supplementary Fig. 5a), and was apparent as early as the third day of co-culture (Fig. 3b). Moreover, this induction required close physical contact between MSCs and cancer cells, because it failed to occur when the two cell populations were separated by a permeable membrane (Fig. 3c).

We undertook to determine the source of the CCL5 produced under conditions of co-culture. To do so, we stably reduced the expression of CCL5 in MDA-MB-231 cells by >80% using short hairpin (sh)RNA (variant siCCL5.1; Supplementary Fig. 6). Importantly, however, subsequent co-culture of these MDA-MB-231.1 cells with MSCs continued to allow accumulation of CCL5 in the culture supernatants to levels that were comparable to those observed in the co-cultures of MSCs and control cancer cells (Fig. 3d). This suggested that the source of CCL5 was the admixed neighbouring MSCs. Indeed, inhibition of CCL5 protein expression in MSCs using the same shRNA hairpin vector (MSC.1; Fig. 3d) resulted in more than 75% reduction of CCL5 protein levels in the co-cultures, indicating that the MSCs were the major source of the CCL5 observed on co-culture of the two cell types. In support of this conclusion, analysis of CCL5 levels in the media of MSCs or MDA-MB-231 cells separated from one another after 3 days of co-culture indicated a strong induction of CCL5 in the culture of MSCs, but not that of BCCs (Supplementary Fig. 5b). Finally, polymerase chain reaction with reverse transcription (RT–PCR) analysis of the RNA prepared from these co-culture-derived MSCs (Supplementary Fig. 5c), as well as from the MSCs isolated from MDA-MB-231+MSC tumours ~4 weeks after tumour implantation (Fig. 3e), indicated a strong accumulation of *CCL5* messenger RNA, suggesting that an active signal transduction pathway is triggered in MSCs by the nearby BCCs.

A series of observations has linked CCL5 signalling and cancer. For example, CCL5 levels in the plasma of breast cancer patients have been correlated with the severity of the disease, and localized CCL5 protein expression was found to be elevated in invasive tumours when compared to *in situ* ductal tumours or benign lesions[16,17]. However, the precise contributions of CCL5 to cancer development and progression are poorly understood. To investigate further the possible causal role of CCL5 in cancer cell metastasis, we overexpressed this chemokine in the MDA-MB-231 BCCs (Supplementary Fig. 7a) and analysed its effects on cancer cell growth and tumorigenesis. The overexpressed CCL5 did not confer any proliferative advantage on cultured cancer cells when compared with those lacking such overexpression (Supplementary Fig. 7b), and had no effect on the ability of BCCs either to grow in an anchorage-independent fashion *in vitro* (Supplementary Fig. 7c), or to form primary subcutaneous tumours in immunocompromised mice (Fig. 4a). However, these tumours exhibited a ~5-fold enhancement in their metastatic potential when compared with control tumours lacking ectopic CCL5 (Fig. 4a). Similarly, overexpression of CCL5 in WI-38 fibroblasts sufficed to enable these cells to promote the metastasis of admixed MDA-MB-231 BCCs (Fig. 4b), indicating that the actions of CCL5 are responsible for much, if not all, of the observed MSC-induced metastasis by the BCCs.

CCL5 promotes lung colonization

Previous reports have described an important role for CCL5 as a chemoattractant for stromal cells, such as macrophages, that express one of the receptors for CCL5, CCR5 (refs 18, 19). Furthermore, CCL5 expression has been associated with increased tumour neovascularization, suggesting that endothelial cells, which express a variety of chemokine receptors, may also be attracted by CCL5 to sites of tumour formation, thereby enhancing tumour angiogenesis[20]. Such observations suggest that CCL5 may contribute to breast cancer metastasis through the recruitment of a number of stromal cell types to sites of primary tumour growth.

However, immunohistochemical analyses indicated that the MDA-MB-231 control and CCL5-overexpressing MDA-MB-231 (MDA-MB-231/CCL5) tumours exhibited comparable numbers of tumour-infiltrating macrophages and had similar vessel densities (as evident by F4/80 and MECA-32 staining for macrophages and endothelial cells, respectively; Supplementary Fig. 8). In addition, we found that ectopic CCL5 expression did not cause an accumulation of other stromal cells, such as SMA-positive cells, in the examined tumours (Supplementary Fig. 8a). Together, these data indicated that the observed CCL5-induced metastasis could not be attributed to significant effects on the numbers of the major constituents of the stroma or to the vascularity of these tumour xenografts.

Invasion and metastatic dissemination of carcinoma cells are often facilitated by their transdifferentiation through the process termed

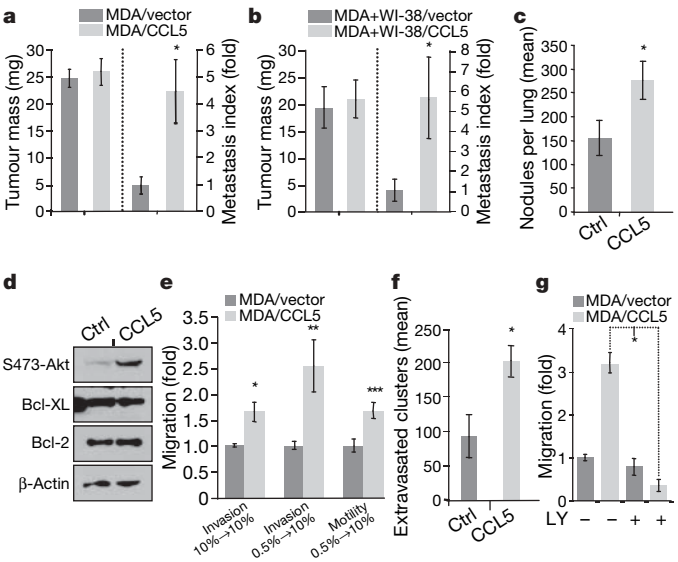


Figure 4 | CCL5 enhances breast cancer cell migration, invasion and metastasis. **a**, A total of 500,000 MDA-MB-231/vector (ctrl) or MDA-MB-231/CCL5 cells were injected subcutaneously in NOD/SCID mice. Tumour masses (mean ± s.e.m., $n = 6$ each group) were taken at 10 weeks. Lung metastasis indices are expressed as fold increase (±s.e.m.) over controls. Data shown are representative of multiple repeats. Asterisk, $P < 0.01$ in one-tailed Student's *t*-test. **b**, A total of 500,000 MDA-MB-231 cells were admixed to 250,000 WI-38 fibroblast controls (WI-38/vector) or WI-38 fibroblasts overexpressing CCL5 (WI-38/CCL5) and were injected subcutaneously in NOD/SCID mice. Tumours ($n = 5$ per group) were excised and weighed at 12 weeks. Masses shown represent mean ± s.e.m. Lung metastasis indices are expressed as fold increase (±s.e.m.) over controls. Asterisk, $P < 0.01$ in one-tailed Student's *t*-test. **c**, A total of 800,000 indicated BCCs were introduced into the circulation of NOD/SCID hosts. GFP-positive cancer colonies in the lungs were counted 6.5 weeks later. Bars represent means ± s.e.m. (MDA-MB-231 controls, $n = 16$ mice; MDA-MB-231/CCL5, $n = 18$ mice). Asterisk, $P < 0.01$ in one-tailed Student's *t*-test. **d**, Western blot analysis of lysates of MDA-MB-231 control or MDA-MB-231/CCL5 cells. β-Actin was used as a loading control. **e**, Transwell migration or Matrigel invasion assays on 50,000 MDA-MB-231 control or MDA-MB-231/CCL5 cells. Data are representative of multiple independent experiments and are expressed as means ± s.d. Asterisk, $P < 0.05$; double asterisk, $P < 0.05$; triple asterisk, $P < 0.01$ in one-tailed Student's *t*-test. **f**, One million GFP-labelled BCCs were injected into the tail vein of NOD/SCID mice. Lungs were processed 48 h later and examined for extravasated cells. Bars represent means ± s.e.m. (MDA-MB-231 cells, $n = 7$ mice; MDA-MB-231/CCL5, $n = 10$ mice). Asterisk, $P < 0.01$ in one-tailed Student's *t*-test. **g**, Transwell migration assays on 50,000 MDA-MB-231 control or MDA-MB-231/CCL5 cells plated with or without the phosphatidylinositol-3-OH kinase inhibitor LY290042 (0.5 μM); representative experiment shown; asterisk, $P < 0.01$ in one-tailed Student's *t*-test.

©2007 Nature Publishing Group

the epithelial-to-mesenchymal transition (EMT), in which cells shed their epithelial characteristics and acquire instead a series of mesenchymal markers that enable their invasiveness and intravasation[21]. Despite their lack of E-cadherin and their expression of detectable levels of mesenchymal markers such as fibronectin (data not shown), the MDA-MB-231 cells studied here exist in an intermediary phenotypic state of 'partial EMT', as they retain a distinctive epithelial morphology *in vitro* and are still responsive to EMT-inducing stimuli in culture. In fact, we observed that ectopic CCL5 expression did not cause MDA-MB-231 cells to undergo the morphological changes usually associated with an EMT (Supplementary Fig. 9a), did not cause rearrangement of their actin cytoskeleton (Supplementary Fig. 9b), and had no impact on the expression of mesenchymal markers closely associated with the EMT process, namely vimentin, N-cadherin (Supplementary Fig. 9c) and fibronectin (data not shown). These data suggested that CCL5 does not directly promote the EMT programme of MDA-MB-231 cells.

We proceeded to explore an alternative possibility: that CCL5 expression affected some of the later, critical steps of the invasion–metastasis cascade, namely the lodging of cancer cells in secondary organs and the subsequent step of colonization. For that purpose, MDA-MB-231/CCL5 cells were injected intravenously into host mice, and the lungs of these hosts were examined ~6 weeks later using fluorescence microscopy. These experiments revealed that CCL5-overexpressing cells indeed had a significant ~1.8-fold advantage over their control counterparts in colonizing the lungs (Fig. 4c), suggesting that CCL5 exposure has effects on later steps of the invasion–metastasis cascade. We note once again that this enhanced tissue-colonizing ability was not due to CCL5's effects on cellular proliferation measured either *in vitro* (Supplementary Fig. 7b) or *in vivo* (Supplementary Fig. 7g, Ki67 staining).

Because improved colonization can be due to enhanced cellular survival, we tested whether CCL5 protects against apoptosis. Notably, we found that MDA-MB-231/CCL5 cells exhibited higher levels of the Ser 473-phosphorylated, activated form of Akt, but exhibited no difference in the levels of other pro-survival proteins, such as Bcl-XL or Bcl-2 (Fig. 4d), or a reduction in the levels of pro-apoptotic molecules such as BAX or BAD (data not shown). Moreover, we found that overexpression of CCL5 had no effect on the ability of MDA-MB-231 cells to withstand serum deprivation (Supplementary Fig. 7b), loss of substrate anchorage (Supplementary Fig. 7d), or hyperoxia (data not shown). We also observed that ectopic CCL5 expression did not protect MDA-MB-231 cells from doxorubicin-induced apoptosis monitored using western blots for cleaved caspase-3 (CC3) and cleaved PARP (as markers of apoptosis; Supplementary Fig. 7e), or TdT-mediated dUTP nick end labelling (TUNEL) assays (Supplementary Fig. 7f). Finally, immunohistochemical analyses on control and CCL5-overexpressing tumours revealed only minor differences in the levels of apoptotic CC3-positive cancer cells among the examined groups (Supplementary Fig. 7g, h). Together, these observations suggested that CCL5 does not exert any detectable pro-survival functions *in vitro* or *in vivo*, and that the observed enhancement of lung colonization was not a consequence of significant anti-apoptotic activities of CCL5.

Akt serves as a key relay switch for upstream signals that promote both cell survival as well as cellular motility[22]. Because CCL5-induced Akt phosphorylation did not correlate with enhanced protection against apoptosis, we tested whether the CCL5-enhanced lung colonization could be due to an increased ability of MDA-MB-231/CCL5 cells to invade from the microvasculature into the lung parenchyma through the process of extravasation. Indeed, ectopic expression of CCL5 enhanced the motility of MDA-MB-231 cells through permeable Boyden chamber membranes by ~1.5-fold as well as the invasion of these cells through Matrigel layers by ~1.6 or ~2.5-fold in either high or low serum conditions, respectively (Fig. 4e). Notably, when we flushed the lungs of mice 48 h after BCC tail-vein injection—in order to remove most cells that remained within the microvasculature of the lungs and thus had not extravasated—we found twice as many deposits in the MDA-MB-231/CCL5-injected group than their control-injected littermates (Fig. 4f). This indicates a clear effect of CCL5 on cancer cell extravasation.

Finally, we investigated the role of Akt in mediating the actions of CCL5 on cellular motility by using the phosphatidylinositol-3-OH

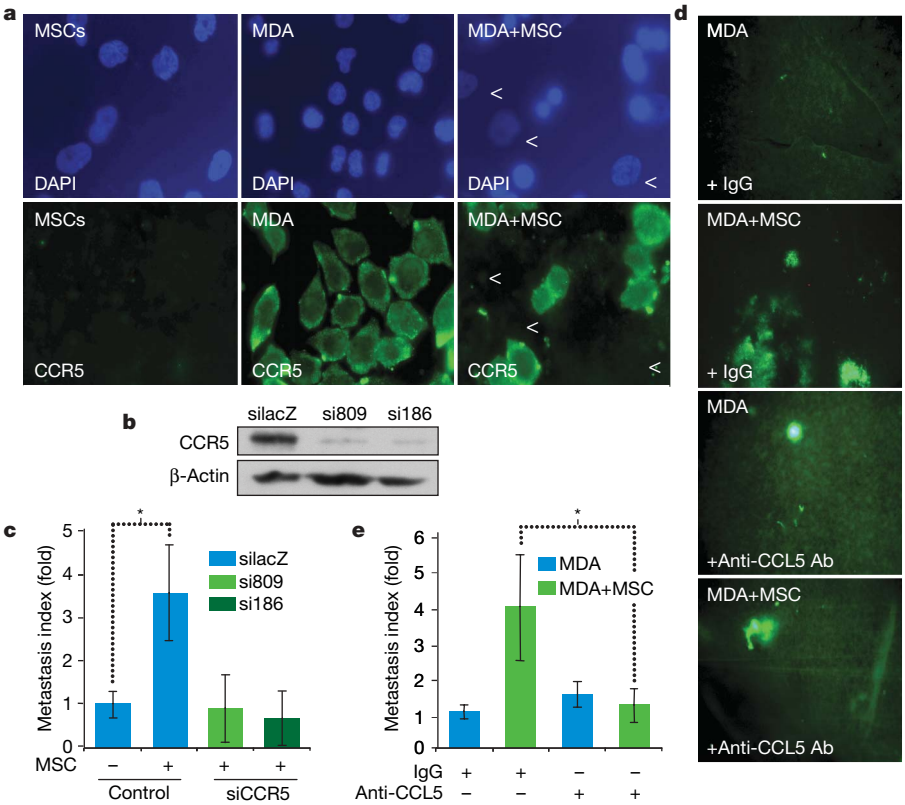


Figure 5 | CCL5–CCR5 interaction is essential for the MSC-induced metastasis. **a**, Immunofluorescence analysis of CCR5 distribution in MDA-MB-231 cells cultured with MSCs. DAPI (for nuclei staining) is in blue; CCR5 detected in green. Arrowheads denote MSCs. **b**, Western blot analysis showing CCR5 expression in MDA-MB-231/silacZ, MDA-MB-231/siCCR5(809) and MDA-MB-231/siCCR5(186) lysates. β-Actin was used as a loading control. **c**, A total of 500,000 cells of the MDA-MB-231 variants in **b** were co-mixed with 1.5×10^6 MSCs and injected subcutaneously into nude mice. Mice were killed when tumours reached 1 cm in diameter and the metastasis index was calculated for each cohort ($n = 5$ per group). Results represent means ± s.e.m.; asterisk, $P < 0.05$ using one-tailed Student's *t*-test. **d**, Anti-CCL5 neutralizing antibody or control IgG was administered intraperitoneally twice weekly in SCID mice bearing MDA-MB-231 ($n = 9$) or MDA-MB-231+MSC tumours ($n = 11$). Representative lung pictures of the indicated cohorts are shown. **e**, Lung metastasis indices of mice in **d**. Data shown are representative of means ± s.e.m. Asterisk, $P < 0.05$ in one-tailed Student's *t*-test.

©2007 Nature Publishing Group

kinase inhibitor LY294002. Drug concentrations that did not inhibit the basal motility levels of MDA-MB-231 cells blocked the elevation of motility induced by ectopic CCL5 expression (Fig. 4g). These results, when taken together, suggest that the observed CCL5-enhanced lung colonization could be ascribed, in significant part, to its ability to promote extravasation and/or motility of cancer cells at sites of dissemination rather than promoting the survival and/or proliferation of these cells.

Essential role for the CCL5–CCR5 loop

CCL5 acts through three G-protein-coupled receptors, termed CCR1, CCR3 and CCR5 (ref. 23). CCR5 has been determined to be the main receptor for CCL5 in MDA-MB-231 cells, as inhibition of its surface expression through dominant-negative mutants abrogated the ability of these cells to respond to CCL5 chemotaxis[24]. We therefore focused our efforts on evaluating the importance of the CCL5–CCR5 interactions in MSC-induced metastasis.

We confirmed that CCR5 is expressed by MDA-MB-231 cells and not by MSCs (Fig. 5a), supporting the notion that MSC-derived CCL5 acts primarily in a paracrine fashion on MDA-MB-231 cells in the BCC and MSC mixed cell populations described above. To probe whether the observed MSC-induced metastasis required CCL5–CCR5 interactions, we inhibited CCR5 expression in MDA-MB-231 cells by more than 85% through shRNA knockdown (ref. 25 and Fig. 5b), and mixed these cells with MSCs before implantation into host mice. Indeed, inhibition of CCR5 expression in the BCCs, achieved using either of two different shRNA constructs, abrogated the ability of MSCs to enhance the metastasis of MDA-MB-231 cells (Fig. 5c). Furthermore, neutralization of CCL5 protein using intraperitoneal injections of an anti-human CCL5 monoclonal antibody also abrogated the MSC-induced metastasis by MDA-MB-231 cells (Fig. 5d, e). In addition, MSCs in which CCL5 expression was inhibited by shRNA knockdown failed to promote metastasis of the admixed MDA-MB-231 cells (data not shown). Taken together, these results underscore the critical importance of the CCL5–CCR5 paracrine interactions in enabling MSCs to induce metastasis of the MDA-MB-231 cells.

Discussion

Certain models of metastatic progression propose that cancer cell invasion and metastasis from the primary tumour site are strongly influenced by contextual signals emanating from the stroma of the primary tumour. It follows that if carcinoma cells are subsequently deprived of such signals, they may revert to an earlier phenotypic state in which they no longer display the traits of high-grade malignancy. Indeed, such a model has been proposed previously by others on the basis of indirect evidence[21]. Here, we demonstrate that at least one mesenchymal cell type, the MSC, can expedite tumour metastasis, and suggest that after primary human carcinomas recruit MSC populations into their midst, subsequent interactions between the MSCs (or their derivatives) and the BCCs endow the latter with invasive and metastatic properties.

Although the recruitment of labelled MSCs to tumour xenografts has been established in a variety of experimental models of tumorigenesis, there is currently no available way to quantify with any accuracy the number of MSCs in actual human tumours, in part because no set of markers has been identified that can uniquely stain these cells without concomitantly staining other mesenchymal types in the tumour-associated stroma[6]. Our demonstration that the stroma derived from tumour xenografts contained appreciable numbers of murine MSCs indicates that significant steady-state levels of these cells are maintained in developing tumours. Interestingly, the use of CD10—one of the markers associated with human MSCs—to purify cells from the stroma of human primary invasive breast carcinomas yielded a population of cells that expresses a number of other markers collectively used to characterize human MSCs (for example, CD44, CD105 and CD106; Fig. 6a). This suggested that, similar to tumour xenografts, human carcinomas also acquire significant numbers of MSCs. Furthermore, we note that CCL5, which is prominent in the stromal gene expression signature associated with poor prognosis of breast cancers[26] (SFT; Fig. 6b, c), is also enriched in the leukocyte- and endothelial cell-free stroma of primary invasive ductal carcinomas (Fig. 6d), specifically in the CD10-positive compartment[27] (Fig. 6e). Collectively, these observations argue strongly for a significant association between stromal CCL5 levels, MSCs and human invasive breast cancers.

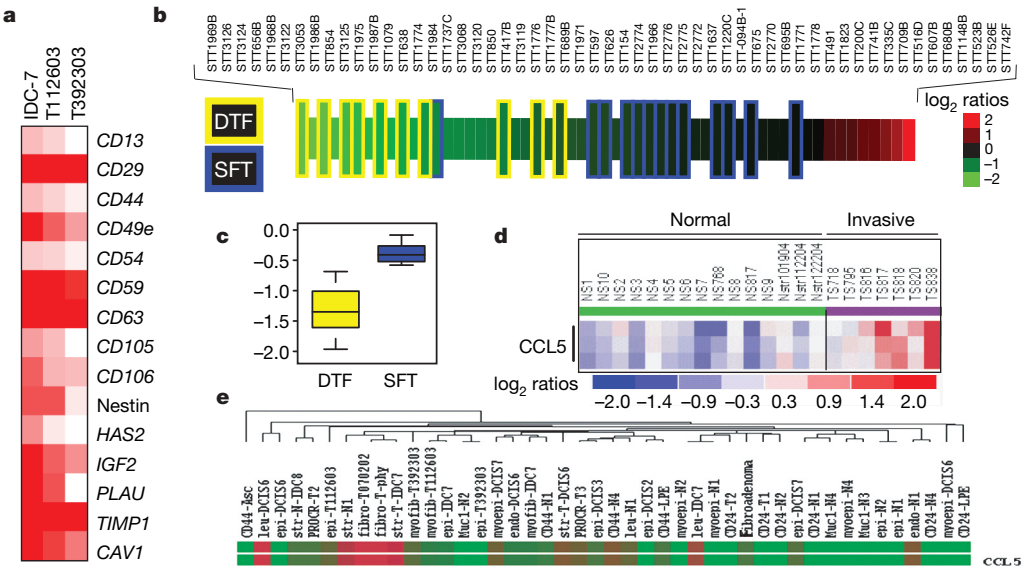


Figure 6 | Stromal fibroblastic cells of human invasive ductal carcinomas are rich in MSC markers and overexpress CCL5. **a**, SAGE TreeView display of MSC markers expressed in stromal CD10-positive cells from invasive tumours[27]. **b**, Soft-tissue tumours were ranked by CCL5 expression[26], from low (green) to high (red). Wide blocks indicate expression ratios of tumours classified as desmoid-type fibromatosis (DTF; yellow outline, $n = 10$) or solitary fibrous tumours (SFT; blue outline, $n = 13$); narrow blocks are other soft-tissue tumours ($n = 32$). **c**, Box plot showing that CCL5 expression is higher ($P = 0.004$) in SFT than in DTF. The difference in $\log_2$ expression ratios between SFT and DTF was tested with the Welch's test. **d**, CCL5 Affymetrix gene expression in the stroma of human invasive ductal cancers compared to that in normal cancer-free breast tissue (indicated as 'Normal'; see Methods). **e**, CCL5 expression is mostly restricted to the CD10-positive fibroblastic cells derived from invasive ductal cancers. The heatmap shown is a cluster of CCL5.genelist obtained as in **a**. Details of the purification methodologies of the various groups indicated in **a**, **d** and **e** are found in ref. 27.

©2007 Nature Publishing Group

Although we have focused here on CCL5 in the MSC–MDA-MB-231 cell interactions, CCL5 seems to have an equally critical involvement in the functional interaction of MSCs with MDA-MB-435 human BCCs. CCL5 levels accumulate synergistically when the two cell types are co-cultured together (Supplementary Fig. 10a), and MSCs in which CCL5 expression was compromised by shRNA knockdown failed to promote metastasis by MDA-MB-435 cells to which they were admixed (Supplementary Fig. 10b). With these facts in mind, we point out that CCL5 does not seem to be involved in regulating the MSC-induced metastasis of MCF7/Ras or HMLER cells, which may depend on other paracrine factors such as VEGF and interleukin-8. Nevertheless, our observations highlight the recently discovered critical roles of chemokine networks in malignant progression[28,29] and suggest the possible utility of a variety of CCL5 analogues and CCR5 antagonists currently used in anti-HIV therapy[30] in treating metastatic disease.

Notably, we have observed that MSCs induce the metastasis of cells to the lung that are, on isolation and re-injection into recipient mice, no more metastatic than their predecessors in the primary tumour (Fig. 2e). This indicated that acquisition of increased metastatic powers by these tumour cells was reversible, and suggested that the maintenance of this phenotype depends on continuing contact with stromal cells. If extended to other tumour types, the present results hold important implications for the molecular analysis of malignant progression. They suggest that many of the cellular functions associated with invasion and metastasis are often not expressed constitutively by carcinoma cells, but rather only transiently in response to contextual signals that tumour cells receive from their stromal microenvironment. If so, analysis of the gene expression patterns of bulk primary tumour populations may fail to detect the expression of key genes mediating invasiveness and metastasis, if only because they are being transiently expressed in minor subpopulations of cells within such tumours. Additionally, attempts at determining the metastatic propensities of tumours may need to be focused on the genes and proteins that confer responsiveness of primary tumour cells to stromal signals, rather than on the genes and proteins that directly mediate the cellular phenotypes of invasion and metastasis.

METHODS SUMMARY

Cells labelled with GFP or ds-red, or harbouring various overexpression or shRNA constructs, were generated by viral transduction followed by FACS enrichment or antibiotic selection. Xenograft experiments were conducted in nude or NOD/SCID mice and metastasis was estimated using fluorescence microscopy. The levels of cytokines, growth factors and chemokines were assessed by immunoassays. Migration and invasion assays were conducted using transwell chambers. Antibody treatment of tumour-bearing mice was conducted by intraperitoneal injections. See Methods for detailed information regarding cell culture, viral infections, *in vivo* colonization and extravasation assays, RT–PCR, TUNEL and anoikis assays, immunohistochemical and immunofluorescence determinations, western blotting, and antibodies used.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 April; accepted 14 August 2007.

1. Bissell, M. J. & Radisky, D. Putting tumours in context. *Nature Rev. Cancer* **1,** 46–54 (2001).
2. Hall, B., Andreeff, M. & Marini, F. The participation of mesenchymal stem cells in tumor stroma formation and their application as targeted-gene delivery vehicles. *Handb. Exp. Pharmacol.* **180,** 263–283 (2007).
3. Pittenger, M. F. *et al.* Multilineage potential of adult human mesenchymal stem cells. *Science* **284,** 143–147 (1999).
4. Young, H. E. *et al.* Clonogenic analysis reveals reserve stem cells in postnatal mammals: I. Pluripotent mesenchymal stem cells. *Anat. Rec.* **263,** 350–360 (2001).
5. Young, H. E. *et al.* Human reserve pluripotent mesenchymal stem cells are present in the connective tissues of skeletal muscle and dermis derived from fetal, adult, and geriatric donors. *Anat. Rec.* **264,** 51–62 (2001).
6. Fox, J. M., Chamberlain, G., Ashton, B. A. & Middleton, J. Recent advances into the understanding of mesenchymal stem cell trafficking. *Br. J. Haematol.* **137,** 491–502 (2007).
7. Gregory, C. A., Prockop, D. J. & Spees, J. L. Non-hematopoietic bone marrow stem cells: molecular control of expansion and differentiation. *Exp. Cell Res.* **306,** 330–335 (2005).
8. Park, C. C., Bissell, M. J. & Barcellos-Hoff, M. H. The influence of the microenvironment on the malignant phenotype. *Mol. Med. Today* **6,** 324–329 (2000).
9. Nakamura, K. *et al.* Antitumor effect of genetically engineered mesenchymal stem cells in a rat glioma model. *Gene Ther.* **11,** 1155–1164 (2004).
10. Nakamizo, A. *et al.* Human bone marrow-derived mesenchymal stem cells in the treatment of gliomas. *Cancer Res.* **65,** 3307–3318 (2005).
11. Hung, S. C. *et al.* Mesenchymal stem cell targeting of microscopic tumors and tumor stroma development monitored by noninvasive *in vivo* positron emission tomography imaging. *Clin. Cancer Res.* **11,** 7749–7756 (2005).
12. Menon, L. G. *et al.* Differential gene expression associated with migration of mesenchymal stem cells to conditioned medium from tumor cells or bone marrow cells. *Stem Cells* **25,** 520–528 (2007).
13. Komarova, S., Kawakami, Y., Stoff-Khalili, M. A., Curiel, D. T. & Pereboeva, L. Mesenchymal progenitor cells as cellular vehicles for delivery of oncolytic adenoviruses. *Mol. Cancer Ther.* **5,** 755–766 (2006).
14. Khakoo, A. Y. *et al.* Human mesenchymal stem cells exert potent antitumorigenic effects in a model of Kaposi's sarcoma. *J. Exp. Med.* **203,** 1235–1247 (2006).
15. Studeny, M. *et al.* Bone marrow-derived mesenchymal stem cells as vehicles for interferon-β delivery into tumors. *Cancer Res.* **62,** 3603–3608 (2002).
16. Luboshits, G. *et al.* Elevated expression of the CC chemokine regulated on activation, normal T cell expressed and secreted (RANTES) in advanced breast carcinoma. *Cancer Res.* **59,** 4681–4687 (1999).
17. Niwa, Y. *et al.* Correlation of tissue and plasma RANTES levels with disease course in patients with breast or cervical cancer. *Clin. Cancer Res.* **7,** 285–289 (2001).
18. Azenshtein, E. *et al.* The CC chemokine RANTES in breast carcinoma progression: regulation of expression and potential mechanisms of promalignant activity. *Cancer Res.* **62,** 1093–1102 (2002).
19. Robinson, S. C. *et al.* A chemokine receptor antagonist inhibits experimental breast tumor growth. *Cancer Res.* **63,** 8360–8365 (2003).
20. Hillyer, P. & Male, D. Expression of chemokines on the surface of different human endothelia. *Immunol. Cell Biol.* **83,** 375–382 (2005).
21. Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nature Rev. Cancer* **2,** 442–454 (2002).
22. Toker, A. & Yoeli-Lerner, M. Akt signaling and cancer: surviving but not moving on. *Cancer Res.* **66,** 3963–3966 (2006).
23. Tanaka, T. *et al.* Chemokines in tumor progression and metastasis. *Cancer Sci.* **96,** 317–322 (2005).
24. Mira, E. *et al.* A role for chemokine receptor transactivation in growth factor signaling. *EMBO Rep.* **2,** 151–156 (2001).
25. Qin, X. F., An, D. S., Chen, I. S. & Baltimore, D. Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. *Proc. Natl Acad. Sci. USA* **100,** 183–188 (2003).
26. West, R. B. *et al.* Determination of stromal signatures in breast carcinoma. *PLoS Biol.* **3,** e187 (2005).
27. Allinen, M. *et al.* Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **6,** 17–32 (2004).
28. Balkwill, F. Cancer and the chemokine network. *Nature Rev. Cancer* **4,** 540–550 (2004).
29. Karnoub, A. E. & Weinberg, R. A. Chemokine networks and breast cancer metastasis. *Breast Dis.* **26,** 75–85 (2006).
30. Palani, A. & Tagat, J. R. Discovery and development of small-molecule chemokine coreceptor CCR5 antagonists. *J. Med. Chem.* **49,** 2851–2857 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Reinhardt for assistance in animal studies, A. Lu for technical help, J. Yao for SAGE data analysis and the MIT Comparative Pathology Laboratory for immunohistochemical analyses. We are grateful to A. Bernad, X.-F. Qin, D. Baltimore and W. Hahn for providing constructs. We would also like to thank R. Hynes, T. Jacks and R. Goldsby for discussions. A.E.K. is a fellow of the Susan G. Komen Breast Cancer Foundation. R.A.W. is an American Cancer Society Research Professor and a Daniel K. Ludwig Cancer Research Professor. This research is supported by grants from the Breast Cancer Research Foundation (R.A.W.), the Ludwig Trust (R.A.W.), the Susan G. Komen Breast Cancer Foundation (R.A.W.) and the Dana-Farber/Harvard Cancer Center Specialized Program of Research Excellence (SPORE) in Breast Cancer (A.E.K., R.A.W. and K.P.).

Author Contributions A.E.K. conceived and designed this study, and performed most experiments; R.A.W. supervised research; A.E.K. and R.A.W. wrote the manuscript; A.B.D. and R.T. provided human MSCs; A.B.D. helped in *in vivo* CCL5 neutralization; A.S. helped in the Luminex screens; A.P.V. and M.W.B. provided technical support in tissue culture, ELISA, western blot, RT–PCR and soft-agar analyses; G.W.B. performed CCL5 analysis on soft tumour expression data; A.L.R. obtained and classified the clinical specimens; K.P. fractionated the clinical samples and performed SAGE analyses; and A.L.R. performed the microarray analysis on sorted stroma.

Author Information The clinical microarray data on the sorted stroma is deposited at http://www.ncbi.nlm.nih.gov/geo, GSE8977. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.A.W. (weinberg@wi.mit.edu).

©2007 Nature Publishing Group

METHODS

Cell lines. The MCF7/Ras[31] and MDA-MB-231 (ATCC HTB-26) cancer cells were infected with pWZL-blasticidin-GFP-expressing retroviral particles and grown in Dulbecco's modified Eagle medium supplemented with 10% calf serum, 100 units ml^{-1} penicillin, 100 $\mu g\,ml^{-1}$ streptomycin and 2 mM L-glutamine (complete medium) at 37 °C in 5% CO_2. The MDA-MB-435 and human mammary epithelial cells HMLER[32] were infected with pRRL3-GFP-expressing lentivirus and grown in complete medium or MEGM media with bovine pituitary extract, respectively. Bone-marrow-derived human MSCs were isolated from hip aspirates of healthy volunteers, propagated as previously described[33], and used between the 4th and 8th passages. Three different batches of MSC cultures derived from three different donors were assayed and exhibited consistent results. MSCs expressing GFP or ds-red were generated by lentiviral (pRRL3-GFP) or retroviral ds-red (Clontech) particles. WI-38 human embryonic lung fibroblasts (ATCC CCL-75) were grown in Dulbecco's modified Eagle medium supplemented with 10% calf serum and were used before the 20th passage. The MDA-MB-231 cells overexpressing human CCL5 (MDA-MB-231/CCL5) were generated by retroviral infection of parental cells with pLZ-CCL5-IRES-gfp[24] viral particles. Control cultures were infected with pLZ-IRES-gfp retrovirus. WI-38 fibroblasts overexpressing CCL5 were generated by retroviral infection of parental cells with pWZL-Blasticidin-CCL5-expressing viral particles. MDA-MB-231 cells lacking CCR5 expression were generated by lentiviral infection of parental cells with FG12-siCCR5(809) or FG12-siCCR5(186) viral particles[25]. MDA-MB-231 cells infected with the control FG12 vector harbouring shRNA against bacterial *lacZ* (FG12-silacZ) were used as a control cell line. All infected cells were enriched for GFP expression using FACS.

Animal studies. All mouse studies were performed under the supervision of MIT's Division of Comparative Medicine and were done in accordance with protocols approved by the Institutional Animal Care and Use Committee. Athymic female nude (NCR nude, nu/nu) mice were purchased from Taconic Laboratories and NOD/SCID mice were bred in-house. Animals were housed under pathogen-free conditions and were given autoclaved food and water *ad libitum*. For xenograft experiments, cancer cells were implanted alone, or were mixed with MSCs or WI-38 fibroblasts and injected subcutaneously into recipient animals as previously described[32]. Nude mice were used at ~10–13 weeks of age and received sub-lethal 400 rad of γ-radiation using a dual ^{137}caesium source 18–24 h before injection. Female NOD/SCID mice were used at 12–14 weeks of age. Tumours were measured twice weekly using precision calipers. Tumour volume was calculated as $4/3\pi r^3$ where r is the estimated radius. Tumours were dissected out at the end of the experiments and weighed.

CFU-F studies. Tumour xenografts were implanted in recipient NOD/SCID females and allowed to grow for 4 weeks. Tumours were then excised, treated with collagenase and the GFP-negative mouse stroma was isolated from the GFP-positive cancer cells using FACS. CFU-F culture assays were performed on the sorted mouse stroma as standard (Stem Cell Technologies). Colonies were stained 14 days later using Giemsa stain and enumerated under light microscopy.

Quantification of lung metastasis. Mice were killed using CO_2 asphyxiation and entire lungs were removed, washed in PBS, and placed in ice-cold Hank's buffer (HBSS, Gibco). Excised lungs were immediately dissected into their various lobes under bright-field microscopy and examined under fluorescence microscopy within 12 h of excision. This circumvents tissue autofluorescence (typically apparent ~24–36 h after necropsy), which greatly masks the GFP signal from the disseminated cancer cells. Lung metastasis burden was estimated as the number of GFP-positive colonies observed under fluorescence microscopy. The lung metastasis index for each mouse was calculated as the ratio of the number of GFP-positive colonies observed in the lungs divided by the mass of the primary tumour (in grams). Indices were pooled within each cohort and were expressed as mean ± s.e.m.

Immunoassays. The levels of cytokines, growth factors and chemokines in the culture media were assessed by xMAP Bio-Plex cytokine array (Bio-Rad Life Sciences) using a Luminex 100 plate reader (Bio-Rad Life Sciences) according to the manufacturer's protocols. The levels of CCL5 in the various culture supernatants were measured using enzyme-linked immunosorbent assay (ELISA; R&D systems).

RT–PCR analysis. GFP-labelled MSCs were admixed to MDA-MB-231 cells and implanted subcutaneously in recipient NOD/SCID mice. Four weeks later, tumours were excised, and GFP-labelled MSCs were purified using FACS and processed for RT–PCR analysis for CCL5 using the following forward and reverse primers: CCL5-left, 5′-TGCAGAGGATCAAGACAGCA-3′, and CCL5-right, 5′-GAGCACTTGCCACTGGTGTA-3′. RT–PCR on cultured cells was performed as standard.

Migration assays. Cancer cells were seeded in the upper well of a 24-well transwell Boyden chamber (8 μm pore size; Costar) and migration was assessed 18 h later. For MSC migration assays, MSCs were layered in the upper well of a 24-well transwell chamber and allowed to migrate towards cell-free media derived from MCF7/Ras or MDA-MB-231 cells placed in the bottom wells. Membranes were processed as standard. Migrating cells were stained with crystal violet and counted using bright-field microscopy.

Extravasation assays. GFP-labelled cancer cells were infused into the circulation of recipient NOD/SCID mice through tail-vein injection. Forty-eight hours later, mice were anaesthetized and their thoracic cage opened. Texas-red lectin (to visualize blood vessels; from *Lycopersicon esculentum*, Vector) was then introduced into the left cardiac ventricle, followed by 4% PFA and then 20 ml of cold PBS. Frozen lung tissue was prepared and sections were processed for fluorescence microscopy as standard.

Anchorage-independent growth assays. We carried out soft agar assays as described previously[34].

Western blot analyses. Western blotting was done using standard protocols. We used primary antibodies against phosphorylated S473-Akt (4051, Cell Signaling), Bcl-XL (2762, Cell Signaling), Bcl-2 (2872, Cell Signaling), β-actin (ab8224, Abcam), CCR5 (ab21653, Abcam), cleaved caspase-3 (9664, Cell Signaling), PARP (9542, Cell Signaling), N-cadherin (205606, Calbiochem) and vimentin (V9, NeoMarkers). We used goat antibodies to mouse (115-035-146) and to rabbit (111-035-144) conjugated with horseradish peroxidase as secondary antibodies (Jackson Immunoresearch), and developed the blots using ECL (Dura, Pierce).

Immunohistochemistry. Immunohistochemical analyses were performed on formalin-fixed, paraffin-embedded tissues. Sections (4-μm thick) were deparaffinized, re-hydrated and subjected to antigen retrieval procedures as described previously[35].

Immunofluorescence. Cells were plated on 0.2% gelatin-coated coverslips in complete medium overnight, washed in PBS, permeabilized in 0.1% Triton-X100, blocked in 1% BSA/10% serum, fixed in 3.6% PBS-buffered paraformaldehyde, and processed for indirect immunofluorescence analyses as standard.

Anti-CCL5 treatment. MDA-MB-231 cancer cells alone or admixed with MSCs were injected subcutaneously into recipient mice. Anti-CCL5 (32 μg per mouse; AF-278-NA, R&D) or control isotype IgG (32 μg per mouse) antibodies were injected into the peritoneum 48 h after tumour implantation and twice weekly for the duration of the experiments (10 weeks).

Anoikis assays. Cancer cells were starved in 1% IFS/DME for 24 h, then trypsinized and suspended in serum-free media. Tubes were continuously rotated at 37 °C for the duration of the experiments. Viable cells were counted using the Trypan blue exclusion assay.

TUNEL assays. Apoptosis quantification analysis was performed using the TUNEL assay as recommended by the manufacturer (Roche), with the exception that the GFP-labelled MDA-MB-231/control and MDA-MB-231/CCL5 cells were fixed for 60 min in ice-cold methanol (instead of 15 min in 3.6% paraformaldehyde) to quench the GFP fluorescence.

Gene expression analysis. Expression ratios (relative to reference mRNA) and classes of soft-tissue tumours were obtained from ref. 26 via NCBI GEO (GSE4305). CCL5 expression was calculated as the mean of probes for IMAGE:1325655 and IMAGE:840753. For Affymetrix gene expression analysis, RNA extraction, cRNA synthesis and hybridization to Human Genome U133 Plus 2.0 Arrays were performed as described previously[36]. Raw expression data obtained using Affymetrix GENECHIP software was normalized and analysed using DNA-Chip Analyser (dChip) custom software (W. H. Wong and C. Li, http://www.dChip.org/). Array probe data were normalized to the mean expression level of each of the three CCL5 probes across stromal samples prepared from a set of 15 normal and 7 IDC (invasive ductal carcinoma) specimens. Leukocyte- and endothelial cell-free stroma was isolated as previously described[27]. Comparisons between 'Normal' and 'Tumour' (IDC breast stroma) were performed using the dChip 'Compare Sample' function. SAGE data were obtained from http://cgap.nci.nih.gov/SAGE/AnatomicViewer and performed as previously described[27]. Data were normalized, log-transformed and clustered using average linkage uncentred analysis. Detailed purification methodologies and sample identification numbers have been previously published[27].

31. Orimo, A. *et al.* Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* **121**, 335–348 (2005).
32. Elenbaas, B. *et al.* Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev.* **15**, 50–65 (2001).
33. Lodie, T. A. *et al.* Systematic analysis of reportedly distinct populations of multipotent bone marrow-derived stem cells reveals a lack of distinction. *Tissue Eng.* **8**, 739–751 (2002).

©2007 Nature Publishing Group

34. Hahn, W. C. *et al.* Creation of human tumour cells with defined genetic elements. *Nature* **400,** 464–468 (1999).
35. Kuperwasser, C. *et al.* Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc. Natl Acad. Sci. USA* **101,** 4966–4971 (2004).
36. Richardson, A. L. *et al.* X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell* **9,** 121–132 (2006).

©2007 Nature Publishing Group

ARTICLES

Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide

Roberto Lande[1], Josh Gregorio[1], Valeria Facchinetti[1], Bithi Chatterjee[3,4], Yi-Hong Wang[1], Bernhard Homey[5], Wei Cao[1], Yui-Hsi Wang[1], Bing Su[1], Frank O. Nestle[6], Tomasz Zal[1], Ira Mellman[3,4], Jens-Michael Schröder[7], Yong-Jun Liu[1] & Michel Gilliet[1,2]

Plasmacytoid dendritic cells (pDCs) sense viral and microbial DNA through endosomal Toll-like receptors to produce type 1 interferons. pDCs do not normally respond to self-DNA, but this restriction seems to break down in human autoimmune disease by an as yet poorly understood mechanism. Here we identify the antimicrobial peptide LL37 (also known as CAMP) as the key factor that mediates pDC activation in psoriasis, a common autoimmune disease of the skin. LL37 converts inert self-DNA into a potent trigger of interferon production by binding the DNA to form aggregated and condensed structures that are delivered to and retained within early endocytic compartments in pDCs to trigger Toll-like receptor 9. Thus, our data uncover a fundamental role of an endogenous antimicrobial peptide in breaking innate tolerance to self-DNA and suggest that this pathway may drive autoimmunity in psoriasis.

Plasmacytoid dendritic cells (pDCs) comprise a dendritic cell population highly specialized for sensing viral and certain microbial infections[1,2]. In contrast to myeloid dendritic cells, pDCs uniquely express Toll-like receptor (TLR)7 and TLR9 (ref. 3), intracellular receptors that recognize viral/microbial nucleic acids within endosomal compartments[4,5]. Together with the constitutive expression of IRF7 (ref. 6) and the ability to retain viral/microbial nucleic acids within endosomal compartments[7], TLR7 and TLR9 permit pDCs to mount rapid and robust type 1 IFN (IFN) responses to viral/microbial infections. At the same time, pDCs normally do not respond to self-DNA[8], which may reflect the fact that viral/bacterial DNA sequences contain multiple CpG nucleotides that bind and activate TLR9, whereas mammalian self-DNA contains fewer such motifs, which are most likely masked by methylation[9,10]. Recent evidence, however, suggests that self-DNA has the potential to trigger TLR9, but may fail to do so because self-DNA somehow fails to access the TLR9-containing intracellular compartments[11,12], although the mechanism of this effect is completely unclear.

On the other hand, it is clear that pDCs do sense and respond to self-DNA in human autoimmune diseases[1]. In systemic lupus erythematosus, pDCs are activated to produce IFNs by circulating immune complexes consisting of autoantibodies and self-nucleic-acids that trigger endosomal TLR following FcγRII-mediated uptake[13–15]. The aberrantly produced IFNs are major effectors in the pathogenesis of autoimmunity[16], mainly by inducing an unabated maturation of peripheral myeloid dendritic cells that stimulate autoreactive T cells[17]. Psoriasis is a common chronic inflammatory disease of the skin in which the local activation of autoimmune T cells induces an abnormal differentiation programme in epidermal keratinocytes[18–20]. We have previously reported that pDCs infiltrate psoriatic skin; their activation to produce IFN represents a key upstream event that initiates the activation of autoimmune T cells, leading to the formation of skin lesions[21]. Physical injury to the skin typically exacerbates psoriasis (the Koebner reaction)[22], and also represents a trigger of pDC activation in psoriatic skin[21], indicating a potential link between skin damage, release of self-DNA and local pDC activation. Yet, how pDCs are activated in psoriasis and what is the molecular mechanism allowing pDCs to sense and respond to self-DNA in human autoimmune diseases are unsolved questions. We now find that LL37, an endogenous antimicrobial peptide overexpressed in psoriatic skin, is the key mediator of pDC activation in psoriasis. LL37 breaks innate tolerance to self-DNA by forming a complex that is delivered to and retained within early endocytic compartments of pDCs to trigger TLR9 and induce IFN production. Thus, we uncover a fundamental mechanism by which pDCs sense and respond to self-DNA coupled with an antimicrobial peptide, and suggest that through this pathway pDCs drive autoimmunity in psoriasis.

LL37 drives pDC activation in psoriasis

To search for a factor that specifically triggers pDCs to produce IFNs in psoriasis, we stimulated peripheral blood pDCs with extracts of psoriatic and healthy skin separated into fractions by preparative reversed-phase HPLC[23]. Whereas extracts of healthy skin were unable to activate pDCs (not shown), psoriatic extracts contained a major IFN-α-inducing fraction, which eluted after 26 min (fraction 26, Fig. 1a). Using electrospray-ionization mass spectrometry (ESI-MS) we identified two principal components of this fraction: a 11,366 Da peptide and a 4,493 Da peptide. The 11,366 Da peptide was psoriasin, as previously reported[24], and the 4,493 Da peptide corresponded to the antimicrobial peptide LL37, as confirmed by sequence analysis after nano-ESI-MS/MS of LysC digests (Supplementary Fig. 1). By using synthetic peptides, we found that only LL37, but neither psoriasin nor GL37, a scrambled version of the LL37 peptide, was able to induce pDC activation or secretion of high levels of IFN-α (Fig. 1b). The concentration of LL37 in cultures containing fraction 26 was 3.9 μM, and the amounts of IFN-α induced were similar to the ones induced by the same concentration of synthetic peptide (Fig. 1c). Furthermore, neutralization of LL37 with a

[1]Department of Immunology, and [2]Department of Melanoma Medical Oncology, M. D. Anderson Cancer Center, University of Texas, Houston, Texas 77030, USA. [3]Department of Cell Biology, Yale University School of Medicine, New Haven, Connecticut 06520-8002, USA. [4]Genentech., 1 DNA Way, South San Francisco, California 94080, USA. [5]Department of Dermatology, Heinrich-Heine-University, Düsseldorf 40225, Germany. [6]St John's Institute of Dermatology, King's College London School of Medicine, London SE1 9RT, UK. [7]Department of Dermatology, University Hospital Schleswig-Holstein, Campus Kiel, University of Kiel, Kiel 24105, Germany.

©2007 Nature Publishing Group

specific antibody completely inhibited the capacity of fraction 26 to induce IFN-α, demonstrating that this activity was entirely mediated by LL37 (Fig. 1c). Anti-LL37 did not affect the control stimulation of pDCs by synthetic TLR7 agonist R837. Thus, we identify in psoriatic skin an endogenous peptide with the unique ability to activate pDCs to produce IFN-α. This peptide is LL37, the only human cathelicidin—a family of cationic antimicrobial peptides induced in damaged skin to provide a rapid broad-spectrum defence against infections by acting as natural antibiotics[25–27].

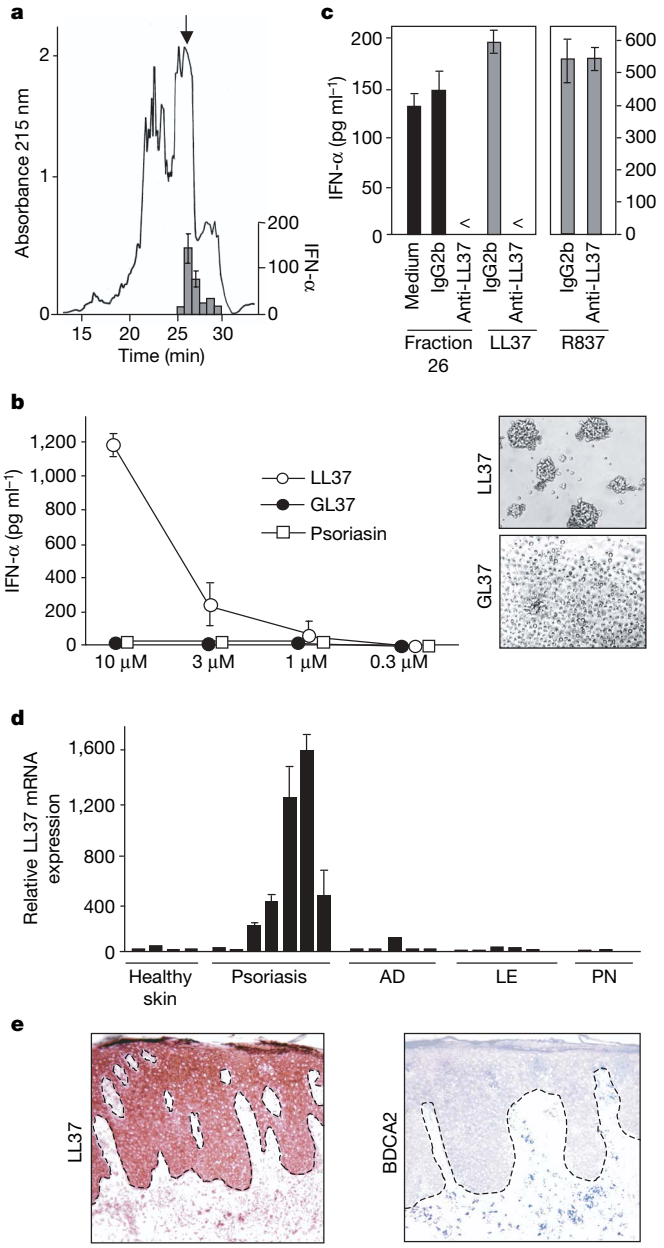


Figure 1 | LL37 is the key trigger of pDC activation in psoriatic skin. **a**, HPLC chromatogram of psoriatic skin extracts. IFN-α produced by pDCs after stimulation with HPLC fractions (inserted bars). Arrow indicates fraction 26. **b**, IFN-α produced by pDCs after stimulation with LL37 (open circles), GL37 (filled circles), or psoriasin (open squares). Data are representative of 10 independent experiments. Right panels show pDC activation with the formation of cell clumps after stimulation with LL37 but not GL37. **c**, IFN-α produced by pDCs after stimulation with fraction 26, LL37 (3.9 μM) or R837 in the presence of anti-LL37 (clone 8A8.2) or control antibodies (IgG2b). <, indicates that the measured value was below the detection limit of the assay (12.5 pg ml^{-1}). **d**, Relative LL37 mRNA tissue expression in healthy skin and lesional skin of psoriasis, atopic dermatitis (AD), lupus erythematosus (LE) and prurigo nodularis (PN). No significant LL37 expression was detected in non-lesional psoriatic skin ($n = 3$, not shown). **e**, Immunohistochemistry for LL37 and pDCs (stained for BDCA2; also known as CLEC4C) in a representative psoriatic skin lesion. **a-d**, Error bars represent the standard deviation of triplicate wells.

We then determined the levels and distribution of LL37 expression in psoriatic skin. We found by real-time PCR analysis that LL37 messenger RNA was highly expressed in psoriatic skin lesions in comparison with the skin of healthy individuals, uninvolved skin of psoriatic patients, or lesions of other common chronic inflammatory skin diseases (Fig. 1d). Immunohistochemistry revealed that LL37 was strongly produced by psoriatic keratinocytes throughout all epidermal layers and was present in the dermal compartment in close association with high numbers of pDCs (Fig. 1e). Because psoriatic skin contains relevant tissue concentrations of LL37 (median concentration of 304 μM)[28], our data strongly suggest that LL37 is the key mediator of pDC activation in psoriasis.

LL37 converts DNA into an activator of pDCs

We next investigated the mechanism by which LL37 triggers pDCs. Because pDCs are specialized sensors of nucleic acids[1,2], and cationic amphipathic peptides with α-helical structure, such as LL37, have been used to transfect DNA[29–32], we asked whether LL37 would activate pDCs in association with self-DNA released into the extracellular environment by dying cells. To test this hypothesis, we depleted DNA from LL37-stimulated pDC cultures by treatment with DNase. We found that DNase strongly abrogated the induction of IFN-α by LL37 but not by the control stimulus R837 (Fig. 2a), indicating that pDC activation by LL37 is mediated by endogenous DNA released into the culture fluids, a finding consistent with the high degree of cell death in pDC cultures. To confirm that LL37 interacts with DNA to trigger pDCs, LL37 was premixed with increasing concentrations of genomic human DNA before stimulation of pDCs. We found that addition of human DNA greatly enhanced the ability of LL37 to induce IFN-α in a dose-dependent manner (Fig. 2b, and Supplementary Fig. 2). In the absence of LL37, human DNA was unable to induce any IFN-α (Fig. 2b), indicating that LL37 induces IFN-α by converting non-stimulatory self-DNA into a potent trigger of pDC activation. Similarly, DNA of bacterial origin became a potent inducer of IFN-α when premixed with LL37 (Supplementary Fig. 2).

Complex formation between LL37 and DNA

To determine whether LL37 directly forms a complex with DNA, we performed size-exclusion HPLC of LL37 premixed with DNA. LL37, but not the control peptide GL37, was able to bind DNA and induce the formation of condensed structures, as illustrated by the reduction in unbound DNA absorbance and the appearance of a broad peak between DNA alone and LL37 alone (Fig. 2c). Binding of LL37 to DNA also induced aggregation of DNA molecules, as shown by the short shoulder eluting before the DNA alone (Fig. 2c). The structural changes in the DNA were readily reversed by salt-induced dissociation of the peptide (Supplementary Fig. 3), indicating an important contribution of ionic interactions between anionic phosphate groups of DNA and cationic amino acids in LL37 during the formation of the complex. Using atomic force microscopy, we visualized LL37 complexed with DNA as prominent nodes of an increased height in the centre of aggregated and coiled DNA (Fig. 2d). The coil-node structures were absent in images of DNA alone (Fig. 2d inset). We conclude that LL37 can form a complex with DNA and induce DNA packaging into aggregated and condensed structures.

DNA coupled with LL37 triggers TLR9

Because pDCs sense DNA through TLR9 (refs 12, 33), we examined whether TLR9 is involved in the recognition of the LL37–DNA complex by pDCs. The IFN-α induced by LL37–DNA was potently inhibited by chloroquine, which blocks endosomal TLR signalling (Fig. 3a). To specifically inhibit TLR9 we used short oligonucleotides (*ODN-TTAGGG*) that block IFN-α induction by a synthetic TLR9

©2007 Nature Publishing Group

agonist (*CpG-2006*) but not the TLR7 agonist (R837) (Fig. 3b). Pretreatment of pDCs with *ODN-TTAGGG* specifically inhibited IFN-α induction by LL37–DNA (Fig. 3b). A TLR9-mediated recognition was confirmed by showing that the LL37–DNA complex activated the *IFNA4* promoter in *IRF-7/TLR-9*-transfected HEK293 cells, but not cells expressing an irrelevant TLR (Fig. 3c). As in pDC cultures, only the LL37–DNA complex but not the DNA alone activated *TLR9*-transfected cells (Fig. 3c). However, in contrast to pDC cultures, LL37 alone was unable to induce activation of *TLR9*-transfected cells, which is consistent with the absence of cell death in these cultures, hence lack of DNA release. Together, these data demonstrate that self-DNA in complex with LL37 can activate pDCs through TLR9.

DNA coupled with LL37 enters pDCs

Intracellular localization of TLR9 in pDCs is thought to be a safeguard against recognition of self-DNA in the extracellular environment while facilitating recognition of viral/microbial DNA released into endosomes during infection[11]. We therefore asked whether the break of innate tolerance to human DNA is the consequence of an LL37-mediated delivery of DNA into the endosomal compartment in pDCs. Human DNA was labelled with the fluorescent dye Alexa 488 and added to pDC cultures. The labelled human DNA was readily detected as associating with pDCs by flow cytometry when given in complex with LL37 but not when delivered alone (Fig. 3d). The complex was internalized, as indicated by the intracellular location of Alexa 488-labelled human DNA and Texas-red-labelled LL37 using confocal microscopy (Fig. 3e, upper panel). The LL37–DNA complexes were contained within membrane-bound compartments, as shown by the colocalization with endosomes stained with the styrene dye FM4-64 (Fig. 3e, lower panel). Importantly, whereas human DNA was undetectable in pDCs when given alone, uncomplexed LL37 readily entered pDCs (Supplementary Fig. 4). Therefore our findings suggest that LL37 enables self-DNA to access the endosomal compartment of pDCs.

LL37 retains DNA in early endosomes

Recent studies[7,34] have suggested that TLR9 signalling in early endosomes leads to IFN-α production by pDCs, whereas signalling TLR9 in late endosomes induces pDC maturation, IL-6 and TNF-α secretion. We found that human DNA in complex with LL37 induces pDC

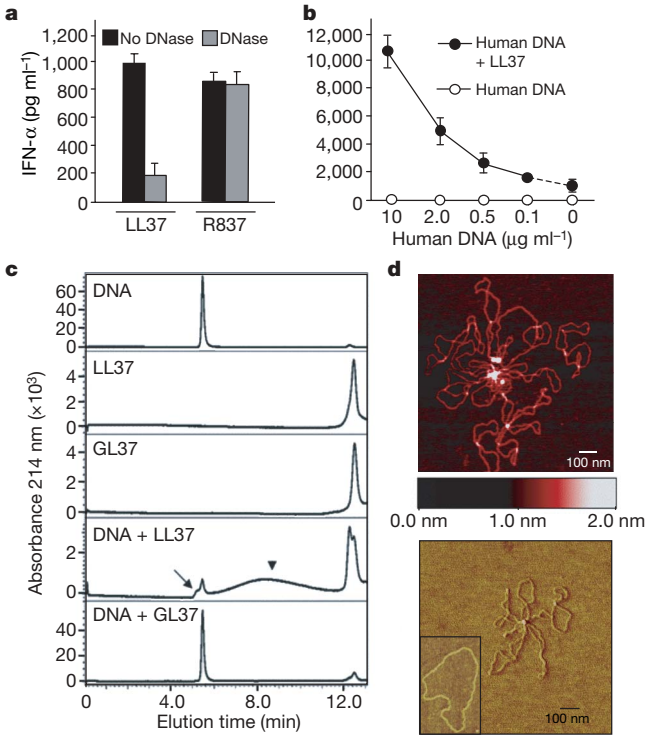


Figure 2 | LL37 converts self-DNA into a potent trigger of pDCs by forming aggregated and condensed structures. **a**, IFN-α produced by pDCs after stimulation with LL37 alone or in the presence of DNase I. Stimulation with R837 was used as a control. **b**, IFN-α produced by pDCs after stimulation with increasing concentrations of human genomic DNA either alone or premixed with LL37. Error bars represent the standard deviation of triplicate wells. Data are representative of at least 5 independent experiments. **c**, Size-exclusion HPLC of LL37 alone, GL37 alone or DNA premixed with LL37 or GL37. The large arrowhead shows DNA condensation, the small arrow shows DNA aggregates. Absorbance scales are different to accommodate the DNA peak. **d**, Atomic force microscopy images of plasmid DNA in complex with LL37. The upper image has been processed in the height mode, a colour bar illustrating the height is given. The lower image has been processed in the phase mode and shows aggregation and coiling of three plasmid DNA molecules around one LL37–DNA node. The insert shows a single plasmid prepared in the absence of LL37.

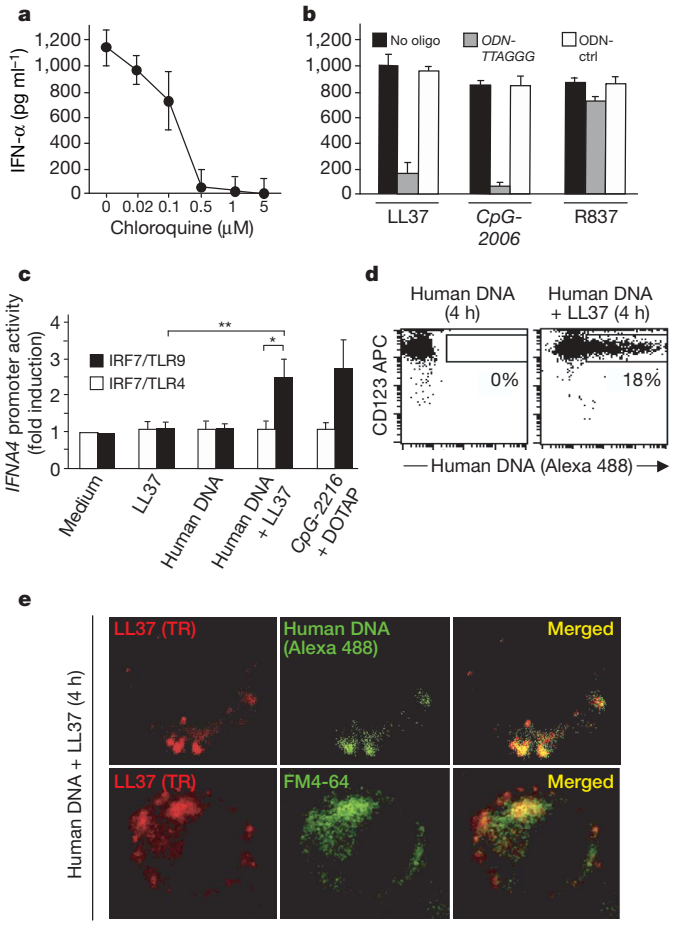


Figure 3 | Self-DNA coupled with LL37 is translocated into pDCs and activates them through TLR9. IFN-α produced by pDCs after stimulation with: **a**, LL37 plus increasing concentrations of chloroquine; or **b**, LL37, *CpG-2006*, or R837 after pretreatment with *ODN-TTAGGG* or control (*ctrl-ODN*). **c**, *IFNA4* promoter activity of *IRF7/TLR9* (solid bars) or *IRF7/TLR4* (empty bars)-transfected HEK293 cells measured by luciferase assay after stimulation at the indicated conditions. *CpG-2216* complexed with cationic lipids (DOTAP) was used as a positive control. **a**, **b**, Data are representative of four independent experiments; error bars represent the standard deviation of triplicate wells. **c**, Data were statistically analysed using unpaired two-tailed Student's *t*-test. Data are presented as mean + s.e.m. of 5 independent experiments; $^{*}P = 0.01$, $^{**}P = 0.003$. **d**, Flow cytometry detection of human DNA^{+} CD123^{+} pDCs stimulated for 4 h with human DNA (Alexa 488) alone or complexed with LL37. **e**, pDCs were stimulated with LL37 (Texas red; TR) in complex with human DNA (Alexa 488) (upper panel), or with LL37 (TR) in complex with unlabelled human DNA plus FM4-64 (lower panel). Cells were imaged by confocal microscopy after 4 h.

©2007 Nature Publishing Group

to produce large quantities of IFN-α, without inducing pDC maturation, IL-6 and TNF-α secretion (Fig. 4a, b), consistent with the possibility that the LL37–human-DNA complex selectively signals by TLR9 in early endosomes. We therefore asked whether LL37 could convert *CpG-2006* (a type B CpG-oligonucleotide) that is thought to stimulate TLR9 in late endosomes, into a ligand that triggers TLR9 in early endosomes[7,34]. Indeed, when complexed with LL37, *CpG-2006* induced pDCs to produce high levels of IFN-α and only low levels of IL-6 and TNF-α as compared with *CpG-2006* alone; the LL37 complex also failed to induce pDC maturation as revealed by the low surface expression of CD80 and CD86 (Fig. 4a, b). To rule out the possibility that this activity resulted from a receptor-mediated effect of LL37 on pDCs, uncomplexed LL37 was given to pDCs either before or after addition of *CpG-2006*. We found that uncomplexed LL37 was unable to enhance IFN production, indicating that the effects of LL37 on TLR9 activation by *CpG-2006* do not occur independently of DNA binding (Supplementary Fig. 5).

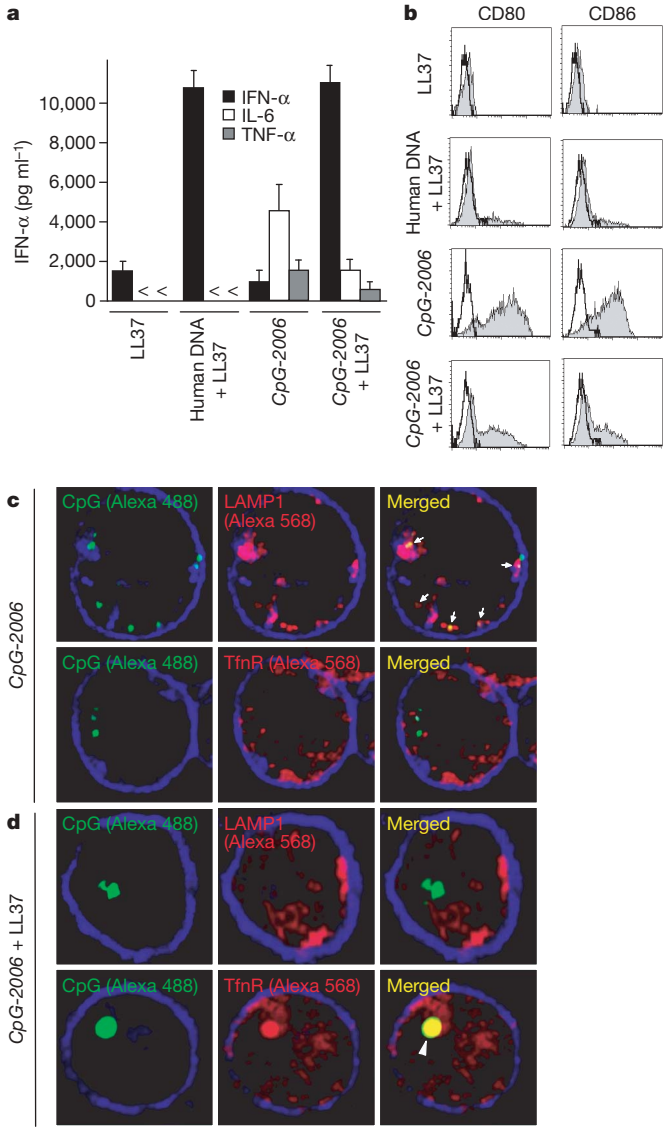


Figure 4 | LL37 retains DNA in early endocytic compartments of pDCs. **a**, IFN-α, IL-6 and TNF-α produced by pDCs after stimulation with LL37 alone, the LL37–human-DNA complex, *CpG-2006* alone, or the LL37–*CpG-2006* complex. **b**, Flow cytometry of stimulated pDCs for CD80 and CD86 surface expression. **a**, **b**, Data are representative of five independent experiments, error bars represent the standard deviation of triplicate wells. **c**, **d**, pDCs were stimulated for 4 h with *CpG-2006* (Alexa 488) alone (**c**), or in complex with LL37 (**d**). Early and late endocytic compartments were stained with Alexa 568-labelled anti-transferrin-receptor (TfnR) and anti-LAMP1, respectively, along with an Alexa 647-conjugated anti-HLA-DR to visualize the contour of the cells. Confocal images of representative cells are shown. Small arrows show co-localization of LAMP1 and *CpG*, and large arrowhead shows co-localization of TfnR and *CpG* in the presence of LL37.

Confocal microscopy revealed that, 4 h after pDC stimulation, uncomplexed *CpG-2006* (Alexa 488) alone partially localized with LAMP1, a marker for late endosomes and lysosomes (Fig. 4c, small arrows), but was distinct from transferrin receptor (TfnR; also known as TFRC), a marker for early and recycling endosomes (Fig. 4c, and Supplementary Fig. 6). In contrast, when complexed with LL37, *CpG-2006* (Alexa 488) appeared as large structures positive for TfnR (Fig. 4d, large arrowheads) and entirely distinct from LAMP1-containing organelles (Fig. 4d, and Supplementary Fig. 6). Thus, the LL37–DNA complex was preferentially retained in a modified early endosomal compartment. These findings were confirmed using *CpG-2216*, a type A CpG-oligonucleotide that induces an early endosomal induction of high levels of IFN-α owing to its ability to spontaneously form aggregated structures. Disruption of *CpG-2216* aggregates into single-stranded structures by heat, and flash cooling abolished its ability to induce IFN-α, however this phenomenon was completely reversed when LL37 was added (Supplementary Fig. 7). Together, these data indicate that the ability of LL37 to convert self-DNA into a trigger of high levels of IFN production depends on its capacity to concentrate and retain DNA in early endocytic compartments, presumably enabling the selective and sustained activation of early endosomal TLR9.

Discussion

Our study identifies the ability of LL37, an endogenous antimicrobial peptide in psoriatic skin, to convert otherwise non-stimulatory self-DNA into a potent trigger of pDCs to produce IFN-α. This process involves three distinct steps. First, LL37 binds DNA to form condensed and aggregated structures, driven by electrostatic interactions. Second, LL37 promotes DNA translocation into the endocytic pathway of pDC, thereby bypassing a safety mechanism for discrimination of viral/microbial from self-nucleic-acids provided by the intracellular localization of TLR9 (ref. 11). Finally, LL37 retains the DNA complex in early endocytic organelles, perhaps modifying them in the process, to specifically sustain the TLR9-mediated IFN response. Thereby, LL37 breaks innate tolerance to self-DNA and, in a similar way to viral responses, elicits a rapid and robust induction of IFNs that initiate innate and adaptive immunity[16]. Among cationic antimicrobial peptides, this function seems to be unique to LL37 because human β-defensin 2, the principal antimicrobial peptide produced in psoriatic skin[35], was unable to induce IFN (data not shown).

Because LL37 is not expressed by healthy skin keratinocytes, but only secreted by these cells at early time points following skin injury[36,37], our results point to a fundamental role of LL37 in alerting the immune system to tissue damage associated with cell death and extracellular release of self-DNA. We propose that through regulated expression of LL37, epithelial cells may actively control the initiation and the duration of immune responses in damaged skin to enhance resistance to infection and potentially initiate wound healing[37]. In psoriasis, LL37 seems to be overexpressed in skin lesions, a finding that has been previously linked to a low rate of infections[28,38]. We now find that LL37 is, in fact, the principal trigger of the pathogenic IFN response in psoriatic skin, providing the first link between an antimicrobial defence system and the pathogenesis of psoriasis. On the basis of our results, we suggest that skin injury coupled with a dysregulated epithelial expression of LL37 may elicit the sustained accumulation and activation of pDCs in psoriatic skin. As a result, the aberrant production of IFNs leads to the maturation of myeloid dendritic cells that initiate local T-cell-mediated autoimmune inflammation and the development of skin lesions. In systemic lupus erythematosus, pDC activation seems to be driven by circulating immune complexes consisting of autoantibody–self-DNA[13–15].

©2007 Nature Publishing Group

Whether LL37 is also involved in pDC activation in systemic lupus erythematosus will be interesting to investigate. Increased expression of LL37 has also been reported in the mucosa of inflammatory bowel disease[39] and the synovia of rheumatoid arthritis[40], suggesting a potential pathogenic relevance of LL37 in other immune-mediated diseases.

In conclusion, this study demonstrates that LL37, an antimicrobial peptide released during skin injury, converts self-DNA into a 'danger signal' that potently activates innate anti-viral-like immune responses. LL37 antagonists may potentially be developed as therapies for psoriasis and other chronic inflammatory diseases, whereas LL37 itself may potentially serve as a vaccine adjuvant.

METHODS SUMMARY

Identification of IFN-inducing factor in psoriasis. Protein extracts of pooled psoriatic scales were fractionated by preparative reversed-phase-HPLC and assayed on purified pDCs for their ability to induce IFN-α. IFN-α-inducing HPLC fractions were analysed by quadrupole-time-of-flight-ESI-MS to identify peptides according to their mass and confirmed by sequencing[24].

Isolation and stimulation of blood pDC. pDCs were isolated from human peripheral blood by MACS-based positive selection for BDCA4 (also known as NRP1) followed by cell-sorting for lineage$^-$CD11c$^-$CD4$^+$ cells by flow cytometry. Purified pDCs were stimulated with LL37 alone, or LL37 premixed with human genomic DNA, or synthetic CpG-oligonucleotides. All reagents are described in Methods. After overnight culture, levels of IFN-α, IL-6 and TNF-α were measured by enzyme-linked immunosorbent assay (ELISA) of the supernatants, and surface expression of CD80 and CD86 was determined by flow cytometry.

Real-time PCR and immunohistochemistry. Were performed as described in the Methods.

Transfection and luciferase reporter assay. HEK293 cells were transfected by lipofection with expression vectors for *TLR9* or *TLR4*, *IRF7* and an *IFNA4* luciferase reporter plasmid. Transfected cells were then stimulated as described and luciferase activity was measured.

Size exclusion HPLC and atomic force microscopy. The DNA–LL37 complex was generated by premixing plasmid DNA with LL37 at a 1:5 mass ratio. Analytical size exclusion chromatography and atomic force microscopy were performed as described in the Methods.

Imaging by confocal microscopy. pDCs were stimulated with human DNA (Alexa 488) or *CpG-2006* (Alexa 488) alone or in complex with LL37. Imaging was performed on living cells or fixed cells using a Leica SP2 RS laser scanning confocal microscope with a water (Leica 63×/1.2 numerical aperture) or an oil immersion objective (63×/1.4 numerical aperture).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 April; accepted 26 July 2007.
Published online 16 September 2007.

1. Liu, Y. J. IPC: professional type 1 interferon-producing cells and plasmacytoid dendritic cell precursors. *Annu. Rev. Immunol.* **23,** 275–306 (2005).
2. Colonna, M., Trinchieri, G. & Liu, Y. J. Plasmacytoid dendritic cells in immunity. *Nature Immunol.* **5,** 1219–1226 (2004).
3. Kadowaki, N. *et al.* Subsets of human dendritic cell precursors express different toll-like receptors and respond to different microbial antigens. *J. Exp. Med.* **194,** 863–869 (2001).
4. Latz, E. *et al.* TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nature Immunol.* **5,** 190–198 (2004).
5. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis e Sousa, C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303,** 1529–1531 (2004).
6. Honda, K. *et al.* IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* **434,** 772–777 (2005).
7. Honda, K. *et al.* Spatiotemporal regulation of MyD88–IRF-7 signalling for robust type-I interferon induction. *Nature* **434,** 1035–1040 (2005).
8. Krieg, A. M. CpG motifs in bacterial DNA and their immune effects. *Annu. Rev. Immunol.* **20,** 709–760 (2002).
9. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124,** 783–801 (2006).
10. Stacey, K. J. *et al.* The molecular basis for the lack of immunostimulatory activity of vertebrate DNA. *J. Immunol.* **170,** 3614–3620 (2003).
11. Barton, G. M., Kagan, J. C. & Medzhitov, R. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nature Immunol.* **7,** 49–56 (2006).
12. Yasuda, K. *et al.* Endosomal translocation of vertebrate DNA activates dendritic cells via TLR9-dependent and -independent pathways. *J. Immunol.* **174,** 6129–6136 (2005).
13. Ronnblom, L., Eloranta, M. L. & Alm, G. V. Role of natural interferon-α producing cells (plasmacytoid dendritic cells) in autoimmunity. *Autoimmunity* **36,** 463–472 (2003).
14. Barrat, F. J. *et al.* Nucleic acids of mammalian origin can act as endogenous ligands for Toll-like receptors and may promote systemic lupus erythematosus. *J. Exp. Med.* **202,** 1131–1139 (2005).
15. Means, T. K. *et al.* Human lupus autoantibody-DNA complexes activate DCs through cooperation of CD32 and TLR9. *J. Clin. Invest.* **115,** 407–417 (2005).
16. Theofilopoulos, A. N., Baccala, R., Beutler, B. & Kono, D. H. Type I interferons (α/β) in immunity and autoimmunity. *Annu. Rev. Immunol.* **23,** 307–336 (2005).
17. Blanco, P., Palucka, A. K., Gill, M., Pascual, V. & Banchereau, J. Induction of dendritic cell differentiation by IFN-α in systemic lupus erythematosus. *Science* **294,** 1540–1543 (2001).
18. Lowes, M. A., Bowcock, A. M. & Krueger, J. G. Pathogenesis and therapy of psoriasis. *Nature* **445,** 866–873 (2007).
19. Griffiths, C. E. & Voorhees, J. J. Psoriasis, T cells and autoimmunity. *J. R. Soc. Med.* **89,** 315–319 (1996).
20. Nickoloff, B. J. & Nestle, F. O. Recent insights into the immunopathogenesis of psoriasis provide new therapeutic opportunities. *J. Clin. Invest.* **113,** 1664–1675 (2004).
21. Nestle, F. O. *et al.* Plasmacytoid predendritic cells initiate psoriasis through interferon-α production. *J. Exp. Med.* **202,** 135–143 (2005).
22. Kupper, T. S. & Fuhlbrigge, R. C. Immune surveillance in the skin: mechanisms and clinical consequences. *Nature Rev. Immunol.* **4,** 211–222 (2004).
23. Harder, J. & Schroder, J. M. Psoriatic scales: a promising source for the isolation of human skin-derived antimicrobial proteins. *J. Leukoc. Biol.* **77,** 476–486 (2005).
24. Glaser, R. *et al.* Antimicrobial psoriasin (S100A7) protects human skin from *Escherichia coli* infection. *Nature Immunol.* **6,** 57–64 (2005).
25. Zanetti, M. Cathelicidins, multifunctional peptides of the innate immunity. *J. Leukoc. Biol.* **75,** 39–48 (2004).
26. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415,** 389–395 (2002).
27. Braff, M. H., Bardan, A., Nizet, V. & Gallo, R. L. Cutaneous defense mechanisms by antimicrobial peptides. *J. Invest. Dermatol.* **125,** 9–13 (2005).
28. Ong, P. Y. *et al.* Endogenous antimicrobial peptides and skin infections in atopic dermatitis. *N. Engl. J. Med.* **347,** 1151–1160 (2002).
29. Dufourcq, J., Neri, W. & Henry-Toulme, N. Molecular assembling of DNA with amphipathic peptides. *FEBS Lett.* **421,** 7–11 (1998).
30. Niidome, T., Wakamatsu, M., Wada, A., Hirayama, T. & Aoyagi, H. Required structure of cationic peptide for oligonucleotide-binding and -delivering into cells. *J. Pept. Sci.* **6,** 271–279 (2000).
31. Kubo, T. & Fujii, M. Specific binding and stabilization of DNA and phosphorothioate DNA by amphiphilic α-helical peptides. *Nucleosides Nucleotides Nucleic Acids* **20,** 1313–1316 (2001).
32. Sandgren, S. *et al.* The human antimicrobial peptide LL-37 transfers extracellular DNA plasmid to the nuclear compartment of mammalian cells via lipid rafts and proteoglycan-dependent endocytosis. *J. Biol. Chem.* **279,** 17951–17956 (2004).
33. Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408,** 740–745 (2000).
34. Guiducci, C. *et al.* Properties regulating the nature of the plasmacytoid dendritic cell response to Toll-like receptor 9 activation. *J. Exp. Med.* **203,** 1999–2008 (2006).
35. Harder, J., Bartels, J., Christophers, E. & Schroder, J. M. A peptide antibiotic from human skin. *Nature* **387,** 861 (1997).
36. Dorschner, R. A. *et al.* Cutaneous injury induces the release of cathelicidin antimicrobial peptides active against group A *Streptococcus*. *J. Invest. Dermatol.* **117,** 91–97 (2001).
37. Heilborn, J. D. *et al.* The cathelicidin anti-microbial peptide LL-37 is involved in re-epithelialization of human skin wounds and is lacking in chronic ulcer epithelium. *J. Invest. Dermatol.* **120,** 379–389 (2003).
38. Frohm, M. *et al.* The expression of the gene coding for the antibacterial peptide LL-37 is induced in human keratinocytes during inflammatory disorders. *J. Biol. Chem.* **272,** 15258–15263 (1997).
39. Schauber, J. *et al.* Heterogeneous expression of human cathelicidin hCAP18/LL-37 in inflammatory bowel diseases. *Eur. J. Gastroenterol. Hepatol.* **18,** 615–621 (2006).
40. Paulsen, F. *et al.* Antimicrobial peptides are expressed and produced in healthy and inflamed human synovial membranes. *J. Pathol.* **198,** 369–377 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Lizee for critical reading of the manuscript; E. Wieder and K. Ramirez from the Flow Cytometry Core Facility for performing cell sorting; J. Wygant from the Monoclonal Antibodies Core Facility for generating the anti-LL37 monoclonal antibody; J. Bartels for ESI-MS analyses; and G. Reyes for technical assistance. This work was funded by a grant from the M. D. Anderson Cancer Foundation to M.G. Part of this work was supported by grants from the 'Deutsche Forschungsgemeinschaft' to J.-M.S. and to B.H.

©2007 Nature Publishing Group

Author Contributions R.L. performed and analysed most of the experiments in this study, and participated in their design. J.G. performed and analysed size-exclusion HPLC and helped in the atomic force microscopy experiments. V.F. and B.S. performed and analysed the experiments with transfected cell lines. B.C. and I.M. helped in establishing and analysing imaging experiments on endosomal trafficking. Yi-Hong Wang performed immunohistochemistry. B.H. collected and characterized clinical samples and participated in the expression analyses. W.C. generated the *IRF7* construct. Yui-Hsi Wang helped in performing RT–PCR experiments. F.O.N. participated along with J.-M.S. and M.G. in the design of experiments with skin extracts. T.Z. helped in performing and evaluating imaging experiments using living cells. J.-M.S. designed, performed and evaluated the separation and biochemical analysis of skin extracts. Y.-J.L. provided critical suggestions and discussions throughout the study. M.G. conceived and supervised this study, was involved in the design and evaluation of all experiments, and wrote the manuscript with comments from co-authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.G. (mgilliet@mdanderson.org).

©2007 Nature Publishing Group

METHODS

Collection of human samples. This study was approved by the Institutional Review Board for human research at the M. D. Anderson Cancer Center in Houston, the Christian-Albrecht University in Kiel, and the Heinrich-Heine University in Düsseldorf. Psoriatic scale material and healthy heel stratum corneum material were collected from multiple donors and pooled as previously described[24]. Skin specimens were taken by 6-mm punch biopsies from normal skin of healthy subject, from uninvolved skin of psoriasis patients and lesional skin of psoriasis, atopic dermatitis, cutaneous lupus erythematosus or prurigo nodularis, defined according to standard clinical and histopathological criteria.

Purification, separation and analysis of skin extracts. Pooled psoriasis scale material from different psoriasis patients or pooled heel stratum corneum from healthy donors were suspended in a cold (4 °C) 1:1 distilled water/ethanol 96% (v/v) mixture acidified to pH 3 by the addition of citric acid and mechanically disrupted using an Ultra-Turrax for 3 h on ice. Subsequently, the suspension was centrifuged for 30 min at 4 °C with 6,000 × g. The particle-free supernatant was concentrated using an Amicon YM-3 membrane (cut-off size 3 kDa) against 10 mM Tris/citrate-buffer, pH 7.4. The concentrated scale extracts were separated by a heparin–sepharose-column (Pharmacia) equilibrated with 10 mM Tris/citrate-buffer, pH 7.4. Peptides were eluted from the column with a continuous gradient of NaCl-solution up to a maximum concentration of 2 M and subjected to a reversed-phase-HPLC using an RP-8-column (Pharmacia) equilibrated with 0.1% trifluoroacetic acid (TFA) in distilled water. The chromatogram was recorded at 215 nm. Elution of peptides was performed with 0.1% TFA/acetonitrile-solution up to 100%. Each fraction (0.5–3 ml) was then lyophilized . Lyophilized peptide fractions were directly dissolved in culture medium and assayed for their ability to induce IFN-α in purified blood pDCs. Fractions with IFN-α-inducing activity were analysed by quadrupole time-of-flight-ESI-MS (Q-TOF 2 mass spectrometer; Waters Micromass) to identify peptides according to their mass. Peptide sequencing was done by collision-induced fragmentation and MS/MS analyses of Lys-C digests with nano-electrospray ionization, as described[24]. Levels of LL37 in the IFN-inducing HPLC fraction were measured using an ELISA kit for LL37 (HyCult Biotechnology).

Reagents. The synthetic peptide LL37 (LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES), and the corresponding scrambled control peptide GL37 (GLKLRFESKIKGEFLKTPEVRFRDIKLKDNRISVQR) were from Innovagen. Synthetic psoriasin was a gift from W. Lu[41]. Synthetic phosphothioate TLR9 agonist *CpG-2006* (1 μM) (5′-tcgtcgttttgtcgttttgtcgtt-3′) was produced by Trilink. Synthetic TLR7 agonist R837 (10 μg ml^{-1}) (imiquimod) was from Invivogen. The specific human TLR9 inhibitor (1 μM) (*ODN-TTAGGG*, 5′-tttagggttagggttagggttaggg-3′) was purchased from Invivogen, the control oligonucleotide sequence (*ctrl-ODN*, 1 μM) (5′-tcctgcaggttaagt-3′) was a gift of F. Barrat (Dynavax). Human genomic DNA (10 μg ml^{-1}) isolated from fetal skin, lung or leucocytes was purchased from BioChain. Endotoxin-free bacterial genomic DNA (10 μg ml^{-1}) isolated from *Escherichia coli* was from Invivogen. Chloroquine was obtained from Sigma-Aldrich, DNase I (5 μg ml^{-1}) was from Boehringer Mannheim. Neutralizing antibody against LL37 was generated by immunizing 6–8-wk-old BALB/c mice with synthetic LL37 peptide. Hybridoma clones secreting monoclonal-antibody specific for LL37 and with the ability to block IFN-induction by LL37 but not unrelated pDC stimuli were established. Neutralizing monoclonal antibody 8A8.2 against LL37 (immunoglobulin 2b) was purified and used at 15 μg ml^{-1}. An IgG2b isotype monoclonal was used as a control.

Isolation of peripheral blood pDCs. For isolation of human peripheral blood pDCs, blood buffy coats of healthy donors were obtained from the Gulf Coast Regional Blood Center, Houston, Texas. After separation of mononuclear cells by Ficoll centrifugation, pDCs were enriched by using anti-BDCA4-conjugated magnetic microbeads (Miltenyi Biotec) and further purified as lineage$^-$CD11c$^-$CD4$^+$ cells using fluorescence-activated cell sorting on a FACSAria (BD Bioscience) to reach 99% purity.

Stimulation of pDC with LL37 or LL37–DNA complexes. Purified pDCs were seeded into 96-well round-bottom plates at 5 × 10^4 cells per well in 200 μl RPMI 1640 (GIBCO) supplemented with 10% FCS (Atlanta biologicals). LL37 (10 μg) was either given directly into pDC cultures (50 μg ml^{-1} = 10 μM), or first premixed with 2 μg of total human DNA (peptide:DNA mass ratio of 5:1) in 10 μl of complete medium. After a 30-min incubation at room temperature the mix was added to the pDC cultures (final concentration was 50 μg ml^{-1} of LL37 and 10 μg ml^{-1} of DNA).

Measurement of cytokine production and maturation of pDCs. Supernatants of stimulated pDCs were collected after overnight culture. Cytokine levels in the supernatants were determined by using ELISA kits for human IFN-α (PBL Biomedical Laboratories), IL-6 and TNF-α (R&D Systems). For analysis of maturation, cultured pDCs were stained with antibodies for CD80 (FITC), and CD86 (PE) (all BD Pharmingen). Flow cytometry was performed on a FACSCalibur (BD Biosciences).

Real-time quantitative PCR. Total RNA from homogenized skin specimen was extracted and reverse-transcribed as previously described[21]. Complementary DNA was quantitatively analysed for the expression of LL37 transcripts by real-time PCR using primers designed against human LL37 (Hs 01011708_m1, Applied Biosystems). Expression was normalized to *GAPDH* mRNA and presented as relative expression as previously described[21]. Expression patterns where confirmed by using 18S rRNA for normalization.

Immunohistochemistry. Immunohistochemical staining was performed on frozen sections of human psoriasis skin specimen[21]. After fixation in acetone, specimen were stained with anti-LL37 (mAb3D11, HyCult Biotechnology) and anti-BDCA-2 (mAbCD303, Miltenyi Biotec) antibodies or isotype control, and the signals were amplified by sequential incubation with biotinylated anti-mouse antibodies and either an avidin peroxidase complex reagent or avidin phosphatase complex reagent (Vector laboratories) before visualization of the immunoreactions.

Luciferase reporter assay. HEK293 cells seeded on 12-well plates (2 × 10^5 cells per well) were transiently transfected by lipofection with 100 ng of *TLR9* or *TLR4* expression vectors (both purchased from Invivogen), 5 ng of *IRF7*, and 100 ng of *IFNA4* luciferase reporter plasmid. Twenty-four hours later cells were collected, replated in 96-well plate and stimulated as indicated for 24 h. *CpG-2216-ODN* complexed with cationic lipids (DOTAP) was used as a positive control for its ability to trigger IFN-α in TLR9/IRF7-expressing cells lines[7]. Luciferase activity was measured with a dual-luciferase reporter assay system (Promega). *Renilla* luciferase reporter gene (pRL-TK) was simultaneously transfected as an internal control for transfection efficiency. *TLR9* and *TLR4* expression vectors were purchased from Invivogen, the *IFNA4* luciferase plasmids were provided by T. Kawai and S. Akira. Full-length human *IRF7* was cloned from cDNA generated from purified human peripheral pDC using a high fidelity Taq polymerase (Invitrogen) with primers acctctagaatggccttggctcctgagagg and attctcgagctaggcgggctgctccagctcc. The *IRF7* gene was then cloned into the pcDNA4 vector (Invitrogen) for transient expression.

Size exclusion HPLC. The LL37–DNA complex was generated by premixing plasmid DNA (pBR322, Invitrogen) with LL37 at 1:5 mass ratio for 30 min. Having a defined length, we used a plasmid DNA for these experiments. Analytical size exclusion chromatography was performed on a Waters Delta 600 HPLC system (Waters) using a TSKgel G4000SW$_{XL}$ column (Tosoh Bioscience). The samples were run at a flow rate of 1.0 ml min^{-1} and UV absorbance was measured at 214 nm. Salt-induced dissociation experiments were done by titrating increasing concentrations of NaCl (0.8 M–3.2 M) on the preformed complex and incubating for 30 min before analysis.

Atomic force microscopy. Atomic force microscopy was performed by Advanced Surface Microscopy. Mica disks, previously cut from blocks of ruby mica, were adhered to glass slides and cleaved using Scotch tape. The cleaved mica surface was then precoated with 3 mM $MgCl_2$ before 25 μl of plasmid DNA (*pBR322*) premixed with LL37 at a 1:5 mass ratio and incubated in phosphate-buffered solution (PBS) for 30 min was added. After a 5 min incubation, the mica surface was rinsed with water and each specimen was placed in the Dimension 3100 large sample AFM (Advanced Surface Microscopy). Tapping mode, capturing height and phase data types for fields ranging from 1 to 10 μm were used. The height data type presents ordinary topographic information and the phase data type provides enhanced edge contrast and the possibility for material contrast. In the height mode, the position of LL37 in complex with DNA was identified as bright tall nodes with a mean height 3.070 nm ± 0.974 ($n = 5$) distinguishable from single-thickness DNA filaments (mean height of 0.378 nm ± 0.176, $n = 5$) or random intersections of DNA filaments (mean height of 1.310 nm ± 0.042, $n = 5$).

Imaging of LL37–human-DNA complexes in pDCs. To visualize the LL37–human-DNA complex in pDCs, human DNA (BioChain) was labelled with Alexa 488 using the Ulysis Nucleic Acid Labelling Kit (Molecular Probes), according to the standard protocol provided by the manufacturer. The method is based on the use of a platinum dye complex that forms a stable adduct with the N7 position of a guanine and, to a lesser extent, adenine bases in DNA. pDCs were stimulated with Alexa 488 human DNA (10 μg ml^{-1}) alone or precomplexed with LL37 (10 μM). After 4 h, cells were extensively washed, stained with APC-labelled anti-CD123 and analysed by flow cytometry. For confocal microscopy pDCs were stimulated with Alexa 488 human DNA (10 μg ml^{-1}) alone or precomplexed Texas-red-labelled LL37 (LL37 (TR), 10 μM; Innovagen). In some experiments, unlabelled human DNA was complexed to Texas-red-labelled LL37 and the green vital dye FM4-64 (20 μM; Molecular Probes) was added to visualize endocytosis[42]. pDCs were washed after 30 min and further incubated for a total of 4 h. Living cells were imaged at 37 °C in complete RPMI supplemented

©2007 Nature Publishing Group

with 10% FCS. Images were acquired by consecutive scanning with the 488 nm line of Argon laser and the 633 nm line of a helium–neon laser. Only one laser line active per scan was used to avoid cross-excitation.

Imaging of endosomal trafficking. pDCs were stimulated with Alexa 488-labelled *CpG-2006* (Trilink) alone or in complex with LL37. After 4 h of stimulation, at least 30% of pDC contained the complex. At this timepoint, cells were washed, transferred onto coverslips pretreated with poly-lysine, fixed for 15 min in 4% para-formaldehyde, and permeabilized in PBS supplemented with 0.1% saponin (Sigma-Aldrich) and 1% normal goat serum (Jackson ImmunoResearch Laboratories). Samples were then stained with Alexa 647-conjugated anti-HLA-DR (BioLegend) together with either biotin-conjugated anti-transferrin-receptor (CD71) or anti-LAMP1 (CD107a) (BD Pharmingen) monoclonal antibodies followed by Alexa 568-labelled streptavidin (Molecular Probes) in permeabilization buffer. Slides were washed and mounted in Prolong Gold antifade mounting media (Molecular Probes) before imaging.

41. Li, X., de Leeuw, E. & Lu, W. Total chemical synthesis of human psoriasin by native chemical ligation. *Biochemistry* **44**, 14688–14694 (2005).
42. Zal, T., Zal, M. A., Lotz, C., Goergen, C. J. & Gascoigne, N. R. Spectral shift of fluorescent dye FM4-64 reveals distinct microenvironment of nuclear envelope in living cells. *Traffic* **7**, 1607–1613 (2006).

©2007 Nature Publishing Group

ARTICLES

Recognition of DNA damage by the Rad4 nucleotide excision repair protein

Jung-Hyun Min[1] & Nikola P. Pavletich[1,2]

Mutations in the nucleotide excision repair (NER) pathway can cause the xeroderma pigmentosum skin cancer predisposition syndrome. NER lesions are limited to one DNA strand, but otherwise they are chemically and structurally diverse, being caused by a wide variety of genotoxic chemicals and ultraviolet radiation. The xeroderma pigmentosum C (XPC) protein has a central role in initiating global-genome NER by recognizing the lesion and recruiting downstream factors. Here we present the crystal structure of the yeast XPC orthologue Rad4 bound to DNA containing a cyclobutane pyrimidine dimer (CPD) lesion. The structure shows that Rad4 inserts a β-hairpin through the DNA duplex, causing the two damaged base pairs to flip out of the double helix. The expelled nucleotides of the undamaged strand are recognized by Rad4, whereas the two CPD-linked nucleotides become disordered. These findings indicate that the lesions recognized by Rad4/XPC thermodynamically destabilize the Watson–Crick double helix in a manner that facilitates the flipping-out of two base pairs.

The autosomal recessive xeroderma pigmentosum syndrome is characterized by sensitivity to sunlight and a highly increased risk of skin cancer (reviewed in ref. 1). There are eight XP complementation groups, seven of which are caused by mutations in genes involved in the nucleotide excision repair (NER) pathway. The lesions that the NER pathway repairs are limited to one DNA strand, but otherwise they are chemically and structurally diverse. They include the pyrimidine–pyrimidine intrastrand crosslinks induced by ultraviolet light, and a wide variety of bulky DNA adducts formed by environmental carcinogens, endogenous metabolites and reactive oxygen species[2].

These lesions are detected either in a genome-wide process, or in a transcription-coupled manner. In global-genome NER, damage recognition requires the XPC protein[3,4], which binds to the lesion and recruits the multi-subunit transcription factor TFIIH. The TFIIH-associated helicases XPB and XPD unwind the damaged DNA, leading to the recruitment of XPA, the single-stranded DNA (ssDNA)-binding replication protein A (RPA), and the nucleases XPG and XPF–ERCC1. This is followed by the excision of a 24–32-nucleotide lesion-containing ssDNA, and by repair synthesis[2,5]. XPC has a central role in the recognition of damage and the initiation of global-genome NER. *In vitro*, XPC can bind specifically to a wide range of NER lesions[6–9], and in ultraviolet-irradiated cells XPC localizes to lesions before the aforementioned NER factors, being strictly required for their subsequent recruitment[10]. Lesion recognition is augmented by the XPB/XPD helicases, which have a proofreading activity required for lesion excision[2,11], and by XPE, which is required for the *in vivo* repair of ultraviolet-induced cyclobutane pyrimidine dimers (CPD) specifically[11]. Recognition of CPDs by XPC is inefficient, and the binding of XPE to CPDs accelerates the localization of XPC to these lesions[11,12].

XPC exists in complex with RAD23B, which participates in several pathways involving ubiquitination and proteasome binding[13]. RAD23B and its orthologues consist of an XPC-binding domain (R4BD), a ubiquitin-like domain (UBL) and two ubiquitin-associated domains (UBA1 and UBA2)[14]. The function of RAD23 in NER is not well understood, but it has been proposed to protect XPC from ubiquitin-mediated degradation[15,16].

NER lesions can distort the structure of the DNA and reduce the stability of the double helix as well. DNA-structure distortion is highly variable. It ranges from adduct intercalation, extrahelical bases and unwinding, which are commonly observed with polycyclic aromatic hydrocarbons and other bulky purine adducts[17–21], to almost regular double helical structures with minor loss of base stacking and Watson–Crick hydrogen bonds, which are typical of the ultraviolet-induced CPD and the 6–4 pyrimidine–pyrimidone photoproduct (6-4PP) lesions[22–24]. In contrast, destabilization of the double helix is ubiquitous across diverse NER lesions and it also correlates well with recognition and repair efficiency[11,25,26]. In fact, multiple base-pair mismatches, which destabilize the DNA duplex but cause only minor structure distortion[27,28], are recognized by XPC comparably to a bona fide NER lesion *in vitro*[29]. Elaborating on this, recent studies have suggested that XPC senses the presence of thermodynamically unstable base pairs on the undamaged strand of the lesion[11,30].

Here we present structures of the *Saccharomyces cerevisiae* orthologue of XPC, known as Rad4, bound to yeast Rad23 and to DNA containing a CPD lesion, and we discuss the implications these structures have for understanding the recognition of diverse NER lesions.

Overall structure of the Rad4–Rad23–DNA complex

Rad4 was truncated to remove 100 amino-terminal and 122 carboxy-terminal residues, which are highly susceptible to limited proteolysis and are presumably unstructured or loosely folded. The truncated Rad4 binds to DNA with an affinity comparable to that of the full-length Rad4 (Supplementary Fig. 1). The crystal structure of the truncated Rad4 bound to a Rad23 fragment containing R4BD and UBA2 domains was determined at 2.35 Å. Co-crystals with damaged DNA were obtained with a CPD lesion embedded within a stretch of three contiguous thymidine–thymidine mismatches (hereafter, CPD-mismatch), previously shown to increase the specificity of XPC binding to lesions[29,30]. The structures of truncated Rad4–Rad23 bound to a 24 base-pair (bp) DNA duplex containing either the CPD-mismatch or the mismatch alone were determined at 3.1 Å and 2.8 Å, respectively (Supplementary Tables 1 and 2).

The Rad4 structure consists of a 310-residue N-terminal α/β domain, followed by three 50–90-residue structurally related α/β

[1]Structural Biology Program and [2]Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.

©2007 Nature Publishing Group

domains characterized by a long β-hairpin (hereafter, BHD1, BHD2 and BHD3, for beta-hairpin domain) (Fig. 1a–c and Supplementary Fig. 2). The N-terminal domain contains the ~45-residue core transglutaminase fold predicted from sequence analysis[31], and we will be referring to it as the TGD (transglutaminase-homology domain) (Supplementary Fig. 3). However, Rad4 lacks the Cys-His-Asp catalytic triad characteristic of the transglutaminase superfamily. The corresponding Rad4 residues have structure-stabilizing roles in a buried environment. The TGD and BHDs pack consecutively, with the TGD forming a large lobe, and the three BHDs extending away from it as a curved protrusion.

Rad4 binds to DNA in two parts: TGD and BHD1 together bind to an 11-bp segment of undamaged, double-stranded DNA (dsDNA; base pairs 1–11), whereas BHD2 and BHD3 together bind to a 4-bp DNA segment that contains the CPD lesion (base pairs 14–17; Fig. 1d). Rad4–DNA contacts extend nearly continuously along these two DNA segments. The Rad4–DNA interfaces bury a total solvent-accessible surface area of 3,450 Å^2, 52% of which is contributed by TGD–BHD1 and 48% by BHD2–BHD3. The Rad23 R4BD and UBA2 domains bind to the Rad4 TGD at a site away from the DNA-binding surface. The UBA2 domain is disordered in the Rad4–Rad23–DNA crystals (Fig. 1 and Supplementary Fig. 4).

BHD2 and BHD3 form a continuous, hand-like structure that interacts with the damaged DNA through two central features. One feature is the BHD3 β-hairpin that inserts through the Watson–Crick double helix, causing the two base pairs that contain the CPD lesion to entirely flip out of the double helix (base pairs 15–16). The other feature is a BHD2–BHD3 groove that holds, like the palm of a hand, the backbone of the undamaged strand and binds to its flipped-out $T15_u$ and $T16_u$ nucleotides ('u' for undamaged strand). The flipped-out CPD-linked $T15_d$ and $T16_d$ ('d' for damaged strand), or corresponding bases of the mismatch-only DNA co-crystal, have no interpretable electron density and we presume they are disordered. They would be in an area fully exposed to solvent and too far from Rad4 to make contacts. There are thus no direct contacts to the lesion. The $T14_u$–$T14_d$ and $G17_u$–$C17_d$ base pairs flanking the two

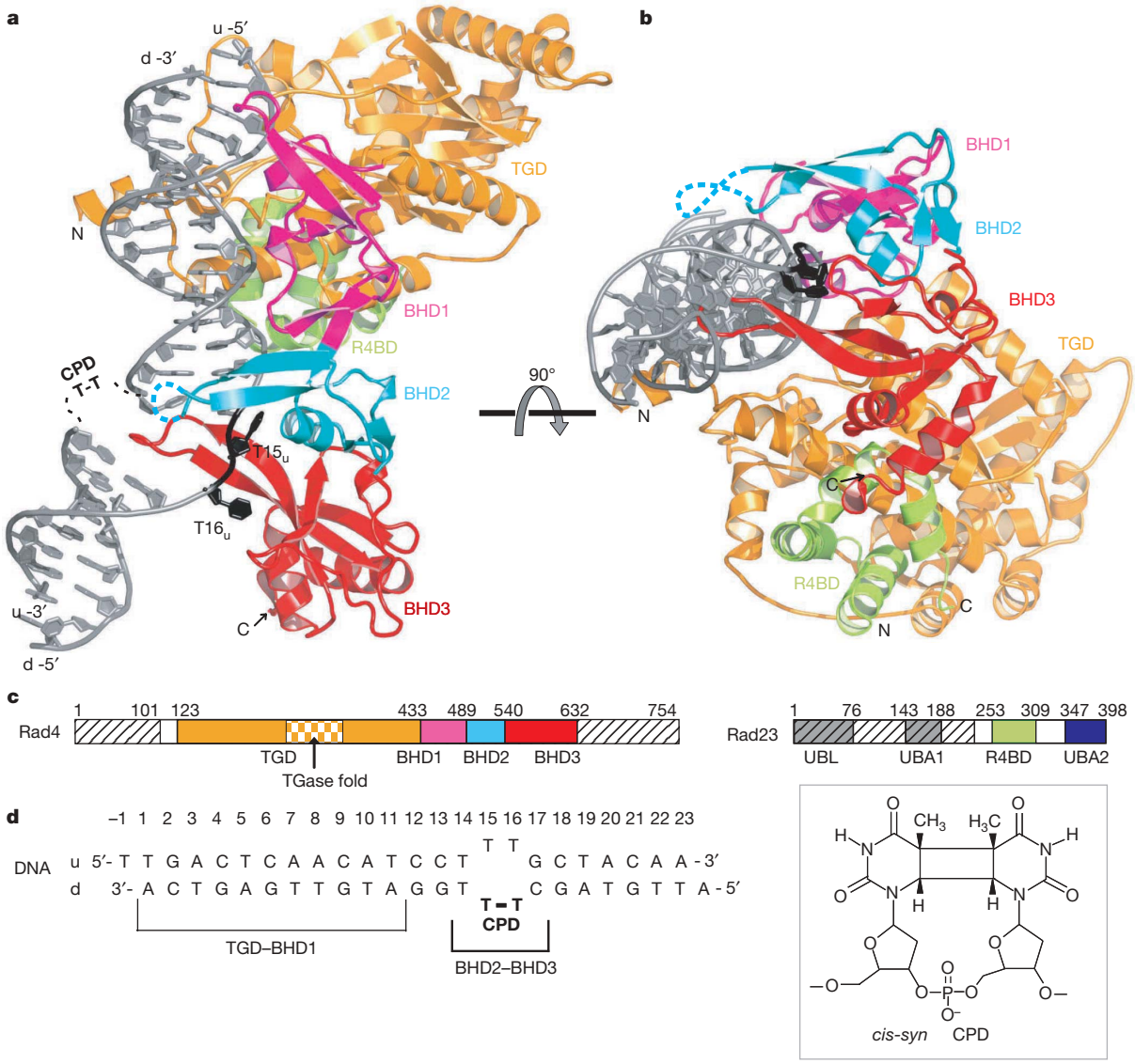


Figure 1 | Overall structures of the Rad4–Rad23–DNA complex. a, Ribbon diagram of the Rad4–Rad23–DNA complex: TGD, gold; BHD1, magenta; BHD2, cyan; BHD3, red; R4BD, green; and DNA, grey. The two disordered CPD-linked thymidines are indicated schematically, and the two flipped-out thymidines bound to Rad4 are coloured black. Dotted lines indicate disordered regions. Figures were prepared with Pymol[43]. **b**, Ribbon diagram of the Rad4–Rad23–DNA complex rotated by 90° about the horizontal axis relative to **a**. The schematic of the disordered CPD is omitted. **c**, Schematic representation of Rad4 and Rad23. Sequences not included in the crystallized proteins are marked with hatched lines. The transglutaminase (TGase) fold is chequered. **d**, Sequence of the CPD-mismatch DNA used in the co-crystallization, and the chemical structure of the CPD lesion. The *cis-syn* isomer describes the relative positions of the pyrimidine rings and the relative orientations of the C5–C6 bonds, respectively.

©2007 Nature Publishing Group

flipped-out base pairs remain in a double helical DNA conformation, with the mismatched $T14_u$–$T14_d$ forming sheared base-pair hydrogen bonds.

The binding of BHD2–BHD3 results in a highly kinked DNA conformation (Supplementary Fig. 5). The double-stranded regions that flank the two flipped-out base pairs are separated by up to ~16 Å, which is ~6 Å farther than the separation they would have in a B-type DNA duplex. This gap lacks a helical twist, and it is associated with a ~42° bend and a ~7 Å translocation of the DNA axis. The DNA conformation and Rad4–DNA contacts are essentially identical in the two Rad4–Rad23–DNA structures.

The structural domains of yeast Rad4 and human XPC share an overall 23% sequence identity (31% for BHD3, 26% for BHD1 and BHD2, and 21% for TGD). Conservation generally maps to residues with roles in stabilizing the structure or in binding to DNA, and in particular to BHD3 residues that contact the damaged DNA segment, indicating that XPC will have the same overall structure as Rad4 and recognize DNA damage in a similar way (Supplementary Fig. 2). On the basis of the Rad4 structure, most non-truncating mutations from xeroderma pigmentosum C patients would be predicted to destabilize the XPC structure.

β-hairpin domains

The three BHDs share a common three-stranded β-sheet (β1, β2 and β3) that supports the long β-hairpin between β2 and β3. The 56-residue BHD1 and 51-residue BHD2 are closely related, having an additional β-sheet (β1′ and β3′) and two short α-helices (α1 and α2) in common (Fig. 2a and Supplementary Fig. 6). BHD1 and BHD2 superimpose with a 1.6 Å root mean square deviation (r.m.s.d.) in the positions of 39 Cα atoms. The larger BHD3 is more divergent, with only 24 of its 91 residues superimposing on BHD2 (1.8 Å, r.m.s.d.). The three BHDs share minimal sequence identity.

The three BHDs pack in an N to C direction and line up along one side of the DNA with an imperfect 26 to 29 Å repeat (Fig. 1a). They are oriented similarly, with their β-hairpins extending towards the DNA, but their interactions with DNA differ substantially. The BHD1 β-hairpin contacts primarily the DNA backbone, whereas the longer BHD3 hairpin inserts through the DNA double helix. The BHD2 hairpin, which is not evolutionarily conserved (Supplementary Fig. 2), makes no DNA contacts and its tip is disordered. The remaining BHD–DNA contacts are made by α0 of BHD1, α1 of BHD2, and a 31-residue BHD3 N-terminal extension that contains an additional β-strand and three β-turns.

The closest structural homologue of the BHDs is the DNA-binding domain of the NER factor XPA (referred to as XPA-MBD-C)[32]. XPA has been shown to bind to ssDNA, to ssDNA–dsDNA junctions, and to denatured bases in cooperation with RPA[33–35]. XPA-MBD-C is more similar to BHD1 and BHD2 (~1.9 Å r.m.s.d. for 31 Cα) than to BHD3 (2.1 Å r.m.s.d. for 22 Cα). Similarities include the three-stranded β-sheet and characteristic β-hairpin, α1 and α2 (Fig. 2b and Supplementary Fig. 6). NMR chemical shift perturbation measurements and mutagenesis have implicated the XPA-MBD-C β-sheet and α1 helix in DNA binding[36,37]. These broadly coincide with the regions of the BHDs involved in DNA binding.

Rad4 interactions with the undamaged dsDNA segment

The TGD structure has a central six-stranded antiparallel β-sheet that packs with an eight-helix subdomain (α1 to α8) preceding it (Supplementary Fig. 3). One end of the sheet is decorated with structural elements that, together with elements from the helical subdomain, form the DNA-binding surface of the TGD.

The TGD and BHD1 form a C-clamp-like structure that binds to the undamaged, double-stranded base pairs 1–11 (Fig. 3a–c and Supplementary Discussion). One tip of the C-clamp, which consists of the TGD α1, α2 and β3–α9 loop, inserts partway into the minor groove, contacting the phosphate groups of the two DNA strands flanking the minor groove (base pairs 5–11). The other tip of the clamp, which consists of the BHD1 α0 and β-hairpin, binds half a turn of DNA away, interacting with the phosphodiester backbone and the flanking major and minor grooves of the DNA.

Overall, 19 TGD–BHD1 side-chain or backbone groups contact 11 consecutive base pairs of dsDNA. Nearly all contacts are made to the phosphodiester and ribose groups of the DNA backbone, and they are roughly equally distributed among the two DNA strands. The interactions of Rad4 with an extended dsDNA segment are consistent with the observations that XPC has substantial affinity for undamaged dsDNA[7,9,29].

Rad4 interactions with the damaged DNA segment

Most of the interactions Rad4 makes with the damaged DNA segment are van der Waals contacts, and they are made primarily to the undamaged strand of the DNA (Fig. 3d, e). The BHD3 β-hairpin approaches the DNA from the major groove and extends through the width of the Watson–Crick double helix. The β-hairpin side

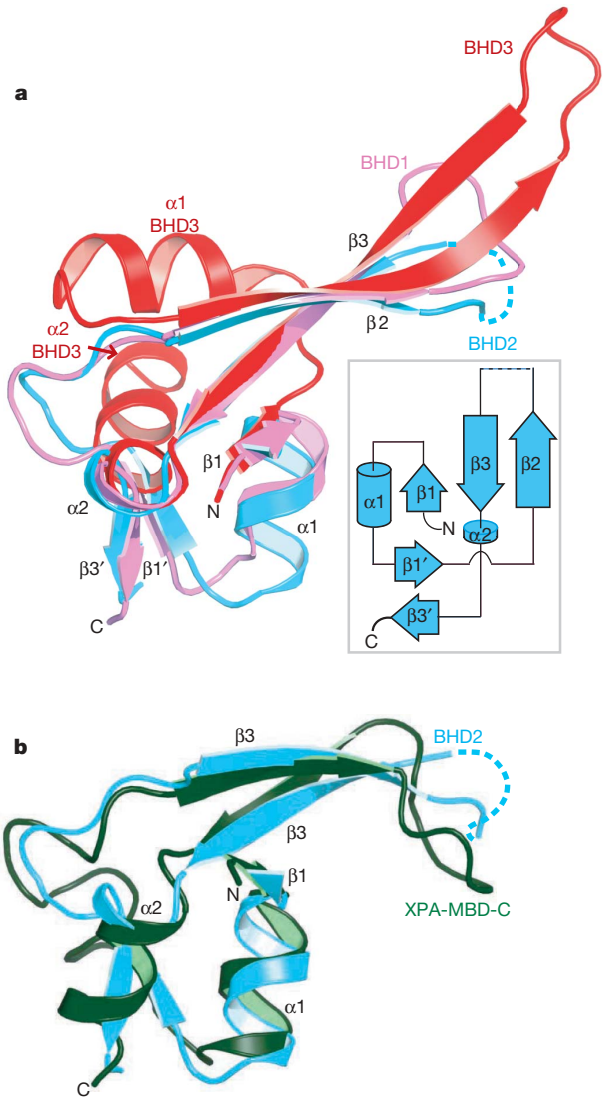


Figure 2 | Rad4 contains three structurally homologous β-hairpin domains. **a**, Superposition of BHD1, BHD2 and BHD3, highlighting the common three-stranded β-sheet and the α1 and α2 helices. The BHD3 N-terminal extension and the BHD1 α0 are omitted for clarity. The topology diagram of BHD2 is shown in an inset. BHD1, pink; BHD2, cyan; and BHD3, red. **b**, The BHDs are structurally homologous to the C-terminal subdomain of the XPA minimal DNA-binding domain (MBD)-C. XPA-MBD-C (green) is superimposed on BHD2 (cyan). The XPA-MBD-C C-terminal helix is omitted for clarity.

©2007 Nature Publishing Group

chains pack with the flanking $T14_u$–$T14_d$ and $G17_u$–$C17_d$ base pairs, with the plane of the hairpin at a steep angle to the base pair planes. Key contacts are made to the flanking bases on the undamaged strand, with the phenyl group of Phe 599 stacking with the $T14_u$ base and the Thr 604 side chain packing with the $G17_u$ base. Contacts to the bases flanking the lesion on the damaged strand are less extensive; they involve aliphatic portions of side-chain and backbone groups from the tip of the hairpin (Glu 600-Arg 601-Gly 602-Ser 603) wedging in between the $T14_d$ and $C17_d$ base and ribose groups.

Of the two flipped-out nucleotides that bind to Rad4, $T16_u$ makes the most extensive contacts. $T16_u$ interacts exclusively with BHD3, with its thymine group being sandwiched between Val 594, Phe 597 and Pro 607 on one face, and Phe 556 on the other, and its ribose group packing with Val 605. There is a single hydrogen bond between the $T16_u$ O4 group and the Asn 554 Nδ2 group. The flipped out $T15_u$ nucleotide interacts with both BHD2 and BHD3. The $T15_u$ ribose group packs with the BHD3 Phe 597 and Phe 599 side chains, whereas the base group makes van der Waals contacts to Met 498 and to the aliphatic portions of Arg 494 and Gln 495 of BHD2. The Arg 494 guanidinium and Gln 495 amide groups in turn hydrogen bond to the phosphate groups of $T16_u$ and $T14_u$, respectively.

The interactions with the flipped-out nucleotides do not show any basis for discriminating among the four bases, consistent with the diversity of NER lesions. The two pockets each have a solvent-exposed side that can accommodate larger purine groups with similar glycosidic bond angles as $T15_u$ and $T16_u$. The single hydrogen bond between Asn 554 and $T16_u$ could, in principle, be made with other bases, or nearby conserved polar residues such as Asn 558 and Glu 560 could provide alternative hydrogen bonds.

The interactions with the flipped-out nucleotides help explain the observations that XPC can bind specifically to artificial substrates such as undamaged ssDNA, short 3′-overhang dsDNA, and DNA with a 3-nucleotide bulge[7,9,38]. All of these DNA structures would mimic aspects of the flipped-out nucleotides of the damaged-DNA conformation in the crystal structure. With 3′-overhang DNA, TGD–BHD1 could also interact with the dsDNA portion, because the 3′-overhang has the same polarity as the flipped-out bases to which BHD2–BHD3 binds. In contrast, XPC shows no binding specificity for short 5′-overhang DNA that has the wrong polarity[38].

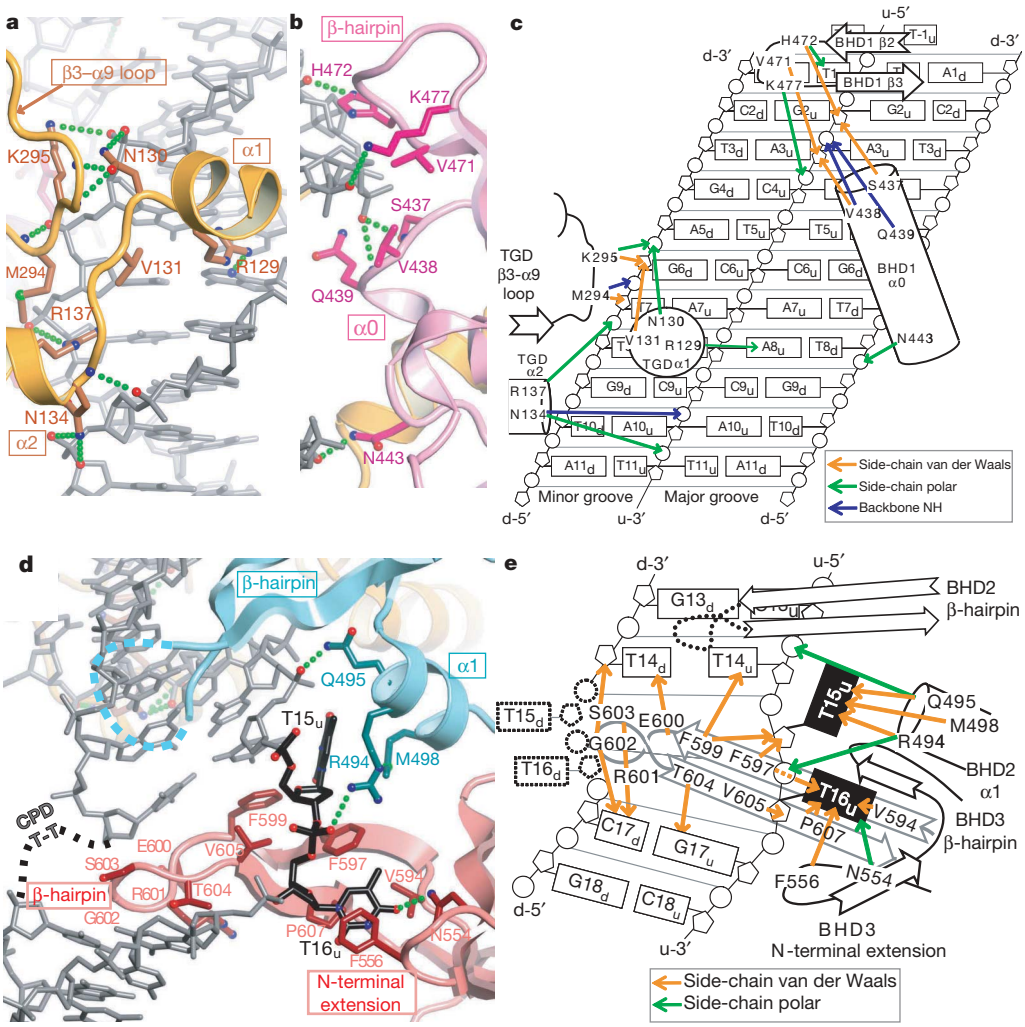


Figure 3 | Rad4 binds to damaged DNA in two parts. **a**, TGD–BHD1 binds to an 11-bp segment of undamaged dsDNA, making extensive contacts to the DNA phosphate and ribose groups. Close-up view of the TGD–DNA interface, in an orientation similar to Fig. 1a, showing side-chain and backbone groups that contact the DNA. Green dotted lines indicate hydrogen bonds. **b**, Close-up view of the BHD1–dsDNA interface, in an orientation slightly rotated about the DNA axis relative to **a**. **c**, Schematic representation of the interactions shown in **a** and **b**. Van der Waals contacts to DNA, orange arrows; polar contacts between side chains and DNA, green arrows; and hydrogen bonds between backbone amide and DNA phosphate groups, blue arrows. **d**, BHD2–BHD3 bind to a 4-bp DNA segment that contains the CPD lesion. Close-up view of the interface in a similar orientation as in Fig. 1a, showing side chains that contact the DNA. The flipped-out thymidines of the undamaged strand are coloured black, and the disordered, CPD-linked thymidines are indicated schematically. **e**, Schematic representation of the interactions shown in **d**. Contacts are marked with arrows, as described in **c**.

©2007 Nature Publishing Group

Rad4 recognizes local destabilization of the DNA duplex

The structure indicates that Rad4/XPC binds to DNA through an induced-fit mechanism, recognizing the propensity of the damaged DNA to adopt the bound conformation of the CPD-mismatch lesion in the crystal structure (open conformation). The adoption of this 'open' conformation will have an energetic cost that is due to the loss of base stacking and Watson–Crick hydrogen bonds. Most NER lesions reduce these energetic costs, as indicated by the lowering of the melting temperature (T_m) of the DNA. Rad4/XPC will form a specific complex if the lesion reduces the energetic cost of adopting the 'open' conformation beyond the binding energy, with the energy difference determining, at least in a thermodynamic sense, the efficiency of recognition. Consistent with this, the 6-4PP lesion causes a larger reduction in the T_m than the CPD[24], and it is recognized by XPC much more efficiently than the CPD[7,29]. On the other hand, certain carcinogenic polycyclic aromatic hydrocarbon adducts do not reduce the T_m, even though they intercalate[39,40], and these lesions evade repair[20].

Induced-fit binding may involve a combination of Rad4/XPC passively sampling the conformational states of the damaged DNA and actively inducing conformational changes. Depending on the lesion, aspects of the 'open' conformation may be well represented in low-energy states, or they may be transiently sampled in higher-energy states. The characteristic of the 'open' conformation that is sensed initially could be flipped-out nucleotide(s), in accord with data suggesting that XPC senses thermodynamically unstable base pairs on the undamaged strand[11,30].

Irrespective of which features are initially sensed, the structure indicates that full binding to two flipped-out nucleotides and complete insertion of the β-hairpin through the double helix will have to occur cooperatively. If BHD2–BHD3 were to bind to the flipped-out nucleotides without opening and kinking the double helix, the bulk of the BHD3 hairpin would clash with the DNA. The structure also suggests that specific binding of Rad4/XPC to a lesion will require the flipping-out of two base pairs. The BHD3 β-hairpin would not fit if only one base pair is flipped out and, conversely, it could not fill a gap resulting from three base pairs flipping out. In fact, the third mismatched base pair ($T14_u$–$T14_d$) of the CPD-mismatch remains in the duplex, adopting sheared base-pair hydrogen bonds. It is thus unlikely that embedding the CPD within a three base-pair mismatch materially changed the bound conformation of the DNA.

Nucleotide flipping is a common feature in enzymes such as DNA glycosylases[41] and DNA methyltransferases[42]. However, most structurally characterized cases involve the flipping-out of a single nucleotide into the enzyme active site, and the extent of DNA opening is much smaller than that observed in the Rad4–Rad23–DNA structure.

Rad4 conformational changes on DNA-binding

Comparison of the apo- and DNA-bound Rad4–Rad23 complexes shows that the individual domains of Rad4 undergo 6–12° rotations over 4–5-residue hinge regions connecting them (Supplementary Fig. 7). When the two structures are aligned on their TGDs, BHD1, BHD2 and BHD3 of the DNA-bound Rad4 are closer to the DNA by 5, 8 and 12 Å, respectively, compared to the apo-Rad4 structure (Fig. 4a). This conformational flexibility in the BHDs helps explain how Rad4/XPC can bind to undamaged dsDNA[7,9,29], and it also suggests a mechanism of lesion detection. In a model of Rad4 bound to undamaged dsDNA, constructed by superimposing a straight B-type DNA duplex on the undamaged dsDNA portion of the damaged DNA co-crystal structure, the lesion-bound conformation of BHD3 would result in extensive clashes with the undamaged dsDNA (Fig. 4b). In contrast, in the apo-Rad4 conformation, BHD3 would not clash with the undamaged dsDNA, but it would be close enough to sample the DNA's conformational space to detect a lesion.

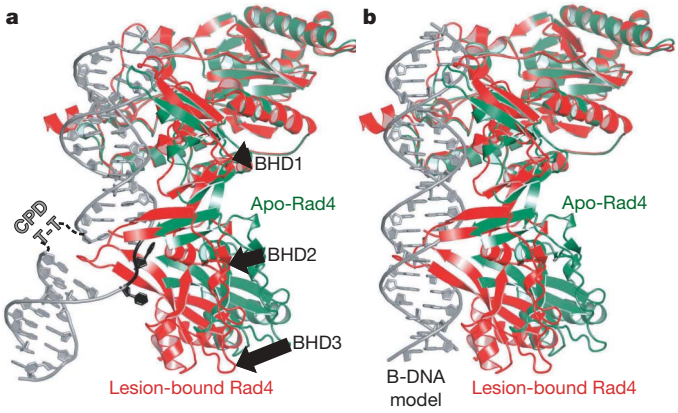


Figure 4 | Rad4 undergoes conformational changes on DNA-binding. **a**, The Rad4–Rad23–DNA complex (red) is superimposed on the apo-Rad4–Rad23 complex (green) by aligning the TGDs. Black arrows indicate the movement of BHD1, BHD2 and BHD3 towards the DNA in the DNA-bound structure relative to the apo-Rad4–Rad23 structure. The BHD3 tip in the apo-Rad4–Rad23 structure is disordered and is not shown. The Rad23 R4BD is omitted for clarity. **b**, Model of undamaged B-type dsDNA bound to Rad4 showing that the apo-Rad4 but not the DNA-bound Rad4 conformation would allow binding to undamaged dsDNA. The model was prepared by superimposing a B-type dsDNA on the undamaged dsDNA portion of the Rad4–Rad23–DNA structure.

Discussion

The structure indicates that Rad4 recognizes lesions that locally destabilize the Watson–Crick double helix in a manner that facilitates the insertion of the β-hairpin through the DNA duplex and the flipping-out of two base pairs. Rad4 does not directly contact the damaged nucleotides, which are disordered in the structure; they would be in a solvent-exposed area too far for Rad4 to contact, irrespective of the chemical nature and size of the lesion. However, because Rad4 binds to the phosphate, ribose and base groups of the two flipped-out nucleotides on the undamaged strand, the lesion would have to be confined to one strand—a requirement for nucleotide excision repair. Our findings thus explain the versatility of global-genome NER in recognizing chemically and structurally diverse lesions, and they provide a structural framework for understanding how certain carcinogen–DNA adducts evade recognition by XPC[20].

METHODS SUMMARY

Briefly, the truncated Rad4–Rad23 complex was overexpressed and purified from insect cells. The structure of the apo-Rad4–Rad23 complex was determined by multiwavelength anomalous dispersion using a $KAu(CN)_2$ derivative. The structures of the Rad4–Rad23–DNA complexes were determined by molecular replacement using the domains of the apo-structure as search models.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 February; accepted 7 August 2007.
Published online 19 September 2007.

1. Cleaver, J. E. Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nature Rev. Cancer* **5**, 564–573 (2005).
2. Gillet, L. C. & Scharer, O. D. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.* **106**, 253–276 (2006).
3. Legerski, R. & Peterson, C. Expression cloning of a human DNA repair gene involved in xeroderma pigmentosum group C. *Nature* **359**, 70–73 (1992).
4. Masutani, C. *et al.* Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J.* **13**, 1831–1843 (1994).
5. Riedl, T., Hanaoka, F. & Egly, J.-M. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* **22**, 5293–5303 (2003).
6. Sugasawa, K. *et al.* Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* **2**, 223–232 (1998).

©2007 Nature Publishing Group

7. Batty, D., Rapic'-Otrin, V., Levine, A. S. & Wood, R. D. Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *J. Mol. Biol.* **300,** 275–290 (2000).
8. Kusumoto, R. *et al.* Diversity of the damage recognition step in the global genomic nucleotide excision repair *in vitro*. *Mutat. Res.* **485,** 219–227 (2001).
9. Hey, T. *et al.* The XPC–HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay. *Biochemistry* **41,** 6583–6587 (2002).
10. Volker, M. *et al.* Sequential assembly of the nucleotide excision repair factors *in vivo*. *Mol. Cell* **8,** 213–224 (2001).
11. Dip, R., Camenisch, U. & Naegeli, H. Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair* **3,** 1409–1423 (2004).
12. Fitch, M. E., Nakajima, S., Yasui, A. & Ford, J. M. *In vivo* recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the *DDB2* gene product. *J. Biol. Chem.* **278,** 46906–46910 (2003).
13. Madura, K. Rad23 and Rpn10: perennial wallflowers join the melee. *Trends Biochem. Sci.* **29,** 637–640 (2004).
14. Masutani, C. *et al.* Identification and characterization of XPC-binding domain of hHR23B. *Mol. Cell. Biol.* **17,** 6915–6923 (1997).
15. Ng, J. M. Y. *et al.* A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes Dev.* **17,** 1630–1645 (2003).
16. Ortolan, T. G., Chen, L., Tongaonkar, P. & Madura, K. Rad23 stabilizes Rad4 from degradation by the Ub/proteasome pathway. *Nucleic Acids Res.* **32,** 6490–6500 (2004).
17. Cosman, M. *et al.* Solution conformation of the (−)-*cis-anti*-benzo[*a*]pyrenyl-dG adduct opposite dC in a DNA duplex: intercalation of the covalently attached BP ring into the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry* **35,** 9850–9863 (1996).
18. Mao, B. *et al.* Solution structure of the (+)-*cis-anti*-benzo[*a*]pyrene-dA ([BP]dA) adduct opposite dT in a DNA duplex. *Biochemistry* **38,** 10831–10842 (1999).
19. de los Santos, C. *et al.* Influence of benzo[*a*]pyrene diol epoxide chirality on solution conformations of DNA covalent adducts: the (−)-*trans-anti*-[BP]G•C adduct structure and comparison with the (+)-*trans-anti*-[BP]G•C enantiomer. *Biochemistry* **31,** 5245–5252 (1992).
20. Geacintov, N. E. *et al.* Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* **65,** 202–210 (2002).
21. O'Handley, S. F. *et al.* Structural characterization of an *N*-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. *Biochemistry* **32,** 2481–2497 (1993).
22. Kim, J. K. & Choi, B. S. The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl(3'→5')thymidine by NMR and relaxation matrix refinement. *Eur. J. Biochem.* **228,** 849–854 (1995).
23. McAteer, K., Jing, Y., Kao, J., Taylor, J. S. & Kennedy, M. A. Solution-state structure of a DNA dodecamer duplex containing a *cis-syn* thymine cyclobutane dimer, the major UV photoproduct of DNA. *J. Mol. Biol.* **282,** 1013–1032 (1998).
24. Jing, Y., Kao, J. F. & Taylor, J. S. Thermodynamic and base-pairing studies of matched and mismatched DNA dodecamer duplexes containing *cis-syn*, (6-4) and Dewar photoproducts of TT. *Nucleic Acids Res.* **26,** 3845–3853 (1998).
25. Gunz, D., Hess, M. T. & Naegeli, H. Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. *J. Biol. Chem.* **271,** 25089–25098 (1996).
26. Buterin, T. *et al.* Unrepaired fjord region polycyclic aromatic hydrocarbon–DNA adducts in *ras* codon 61 mutational hot spots. *Cancer Res.* **60,** 1849–1856 (2000).
27. Gao, Y. G., Robinson, H., Sanishvili, R., Joachimiak, A. & Wang, A. H. Structure and recognition of sheared tandem G•A base pairs associated with human centromere DNA sequence at atomic resolution. *Biochemistry* **38,** 16452–16460 (1999).
28. Chou, S. H. & Chin, K. H. Solution structure of a DNA double helix incorporating four consecutive non-Watson–Crick base-pairs. *J. Mol. Biol.* **312,** 769–781 (2001).
29. Sugasawa, K. *et al.* A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev.* **15,** 507–521 (2001).
30. Buterin, T., Meyer, C., Giese, B. & Naegeli, H. DNA quality control by conformational readout on the undamaged strand of the double helix. *Chem. Biol.* **12,** 913–922 (2005).
31. Anantharaman, V., Koonin, E. V. & Aravind, L. Peptide-N-glycanases and DNA repair proteins, Xp-C/Rad4, are, respectively, active and inactivated enzymes sharing a common transglutaminase fold. *Hum. Mol. Genet.* **10,** 1627–1630 (2001).
32. Ikegami, T. *et al.* Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nature Struct. Biol.* **5,** 701–706 (1998).
33. Buschta-Hedayat, N., Buterin, T., Hess, M. T., Missura, M. & Naegeli, H. Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA. *Proc. Natl Acad. Sci. USA* **96,** 6090–6095 (1999).
34. Yang, Z. *et al.* Specific and efficient binding of xeroderma pigmentosum complementation group A to double-strand/single-strand DNA junctions with 3'- and/or 5'-ssDNA Branches. *Biochemistry* **45,** 15921–15930 (2006).
35. Wang, M., Mahrenholz, A. & Lee, S. H. RPA stabilizes the XPA-damaged DNA complex through protein–protein interaction. *Biochemistry* **39,** 6433–6439 (2000).
36. Buchko, G. W. *et al.* DNA–XPA interactions: a ^{31}P NMR and molecular modeling study of dCCAATAACC association with the minimal DNA-binding domain (M98–F219) of the nucleotide excision repair protein XPA. *Nucleic Acids Res.* **29,** 2635–2643 (2001).
37. Camenisch, U., Dip, R., Schumacher, S. B., Schuler, B. & Naegeli, H. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nature Struct. Mol. Biol.* **13,** 278–284 (2006).
38. Sugasawa, K., Shimizu, Y., Iwai, S. & Hanaoka, F. A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair* **1,** 95–107 (2002).
39. Cosman, M. *et al.* Solution conformation of the (+)-*trans-anti*-[BPh]dA adduct opposite dT in a DNA duplex: intercalation of the covalently attached benzo[*c*]phenanthrene to the 5'-side of the adduct site without disruption of the modified base pair. *Biochemistry* **32,** 12488–12497 (1993).
40. Cosman, M. *et al.* Solution conformation of the (−)-*trans-anti*-benzo[*c*]phenanthrene-dA ([BPh]dA) adduct opposite dT in a DNA duplex: intercalation of the covalently attached benzo[*c*]phenanthrenyl ring to the 3'-side of the adduct site and comparison with the (+)-*trans-anti*-[BPh]dA opposite dT stereoisomer. *Biochemistry* **34,** 1295–1307 (1995).
41. Huffman, J. L., Sundheim, O. & Tainer, J. A. DNA base damage recognition and removal: new twists and grooves. *Mutat. Res.* **577,** 55–76 (2005).
42. Cheng, X. & Blumenthal, R. M. Finding a basis for flipping bases. *Structure* **4,** 639–645 (1996).
43. DeLano, W. L. The PyMOL Molecular Graphics System ⟨http://www.pymol.org⟩ (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. King for mass spectroscopic analysis; H. Erdjument-Bromage for N-terminal sequencing; the staff of the Advanced Photon Source ID-24 and 8-BM beamlines for help with data collection; M. Minto for administrative assistance; and Y. Goldgur, A. Wong, A. Smalls-Mantey, A. Rozenbaum and the members of the Pavletich laboratory for help and discussions. This work was supported by the NIH and the Howard Hughes Medical Institute. J.-H.M. was supported by the Leukemia & Lymphoma Society as a Special Fellow.

Author Information Coordinates and structure factors of the Rad4–Rad23 and Rad4–Rad23–DNA complexes have been deposited in the Protein Data Bank under accession code 2QSF (Rad4–Rad23), 2QSG (Rad4–Rad23 bound to CPD-mismatch DNA) and 2QSH (Rad4–Rad23 bound to mismatch-only DNA). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.P.P. (pavletin@mkscc.org).

©2007 Nature Publishing Group

METHODS

Protein expression and purification. Limited proteolysis of the insect-cell-expressed 754-residue Rad4 showed that the 100 N-terminal (residues 1–100) and 122 C-terminal (633–754) residues are highly susceptible to digestion, suggesting that they are unstructured or loosely folded. Consistent with this, the two regions have overall low sequence complexity and are not conserved, except for a 23-residue segment (645–667) that is thought to act as a linear element[44,45] in the middle of the C-terminal TFIIH-binding region.

The optimal production of soluble Rad4 and human XPC required co-expression of them with their respective Rad23 partner. Testing of various deletion constructs of the 398-residue yeast Rad23 showed that the R4BD, UBA2 and UBL domains were required for maximal Rad4 recovery from cell extracts, whereas UBA1 was dispensable. UBL was also dispensable after the Rad4–Rad23 complex was purified.

To reduce the structural flexibility of the Rad4 and Rad23 proteins and increase the probability of crystallization, an expression vector (pFastBac Dual; Invitrogen) was constructed to produce an N-terminal His_6-tagged Rad4 protein (101–632) together with a Rad23 fragment in which the UBA1 domain (135–229) was replaced with a thrombin protease site.

The Hi5 insect cells co-expressing the Rad4–Rad23 complex were harvested 2 days after infection and were lysed in 25 mM Tris-HCl, 500 mM NaCl, 0.5 mM tris(2-carboxyethyl) phosphine hydrochloride, 0.5 mM Pefabloc SC (Roche), 1 $\mu g\,ml^{-1}$ each of leupeptin, aprotinin and pepstatin, pH 7.5, at 4 °C, using a cell homogenizer (Avestin). The proteins were first purified using His-Select Nickel agarose resin (Sigma). They were then dialysed against 25 mM Tris-HCl, 50 mM NaCl, 4 mM dithiothreitol (DTT) and 2 mM ethylenediamine tetra-acetate, pH 8.8, and were purified by anion exchange chromatography (Source Q, Amersham), followed by thrombin digestion of the His_6 tag and of the UBL domain (1–134). The complex was further purified by cation exchange (Source S, Amersham) and gel-filtration (Superdex200, Amersham) chromatography, and it was concentrated by ultrafiltration to $\sim$10 $mg\,ml^{-1}$ in 5 mM bis-tris propane–HCl (BTP-HCl), 400–800 mM NaCl, 5 mM DTT, pH 6.8. The final Rad4–Rad23 complex used in crystal-structure determinations contains residues 101–632 of Rad4, and residues 230–398 of Rad23.

Crystallization and data collection. All crystals were grown by the hanging-drop vapour diffusion method at 4 °C. The apo-Rad4–Rad23 crystals were grown from 100 mM sodium phosphate, 8% (w/v) polyethylene glycol (PEG) 2000, 2 mM DTT, pH 6.5, and were harvested in crystallization buffer containing 11% (w/v) PEG 2000. Heavy atom soaks were performed in harvest buffer lacking DTT with 3 mM $KAu(CN)_2$ for 12 h. Crystals were flash-frozen in harvest buffer supplemented with 25% (v/v) ethylene glycol. Rad4–Rad23–DNA ternary complexes were prepared for crystallization by mixing the purified Rad4–Rad23 complex with a 1.1 molar ratio of DNA, followed by dialysis against 5 mM BTP-HCl, 200 mM NaCl, 10 mM DTT, pH 6.8. The Rad4–Rad23–DNA co-crystals were typically grown from 50 mM BTP-HCl, 100 mM NaCl, 6–15% (v/v) isopropanol, 10–14 mM calcium chloride, 5 mM DTT, pH 6.8. They were harvested in crystallization buffer containing 8–18% (v/v) isopropanol, and they were flash-frozen in harvest buffer supplemented with 20% (v/v) PEG 400. Diffraction data were collected at −170 °C and were processed with the HKL2000 suite[46].

Structure determination and refinement. The structure of the Rad4–Rad23 complex was determined by the multiwavelength anomalous dispersion (MAD) method using the $KAu(CN)_2$ derivative. Phases were calculated with the program SHARP (Global Phasing) followed by density modification with SOLOMON[47]. The model was built with the program O[48] and was improved by several cycles of manual rebuilding and refinement with CNS[49] or REFMAC[50]. The refined model contains residues 126–300, 304–515, 528–599 and 606–632 of Rad4, residues 255–309 and 347–395 of Rad23, and 250 water molecules. The structures of the Rad4–Rad23–DNA complex were determined by molecular replacement with AMORE[50], using the structures of the Rad4 TGD–BHD1 and BHD2–BHD3 as search models. The final models of the ternary complexes contain residues 123–517 and 526–632 of Rad4, and 256–369 of Rad23.

44. Yang, A. *et al.* Flexibility and plasticity of human centrin 2 binding to the xeroderma pigmentosum group C protein (XPC) from nuclear excision repair. *Biochemistry* **45,** 3653–3663 (2006).
45. Thompson, J. R., Ryan, Z. C., Salisbury, J. L. & Kumar, R. The structure of the human centrin 2-xeroderma pigmentosum group C protein complex. *J. Biol. Chem.* **281,** 18746–18752 (2006).
46. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
47. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276,** 472–494 (1997).
48. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
49. Brunger, A. T. *et al.* Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
50. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).

©2007 Nature Publishing Group

Vol 449|4 October 2007|doi:10.1038/nature06210

LETTERS

Extremely fast acceleration of cosmic rays in a supernova remnant

Yasunobu Uchiyama[1], Felix A. Aharonian[2,3], Takaaki Tanaka[1,4], Tadayuki Takahashi[1] & Yoshitomo Maeda[1]

Galactic cosmic rays (CRs) are widely believed to be accelerated by shock waves associated with the expansion of supernova ejecta into the interstellar medium[1]. A key issue in this long-standing conjecture is a theoretical prediction that the interstellar magnetic field can be substantially amplified at the shock of a young supernova remnant (SNR) through magnetohydrodynamic waves generated by cosmic rays[2,3]. Here we report a discovery of the brightening and decay of X-ray hot spots in the shell of the SNR RX J1713.7−3946 on a one-year timescale. This rapid variability shows that the X-rays are produced by ultrarelativistic electrons through a synchrotron process and that electron acceleration does indeed take place in a strongly magnetized environment, indicating amplification of the magnetic field by a factor of more than 100. The X-ray variability also implies that we have witnessed the ongoing shock-acceleration of electrons in real time. Independently, broadband X-ray spectrometric measurements[4] of RX J1713.7−3946 indicate that electron acceleration proceeds in the most effective ('Bohm-diffusion') regime. Taken together, these two results provide a strong argument for acceleration of protons and nuclei to energies of 1 PeV (10^{15} eV) and beyond in young supernova remnants.

RX J1713.7−3946 (ref. 5) is a unique remnant of a supernova in the sense that its X-ray emission is strongly dominated by a non-thermal component[6–9], which has been presumed to be synchrotron radiation of ultrarelativistic electrons[10,11]. Within the generally accepted theory of diffusive shock acceleration (DSA; reviewed in refs 12, 13), the extension of synchrotron radiation into the X-ray domain requires a high shock speed of about 3,000 km s^{-1} (refs 6, 14). The sub-arcsecond angular resolution of NASA's Chandra X-ray observatory permits the measurement of the shock speed of young, relatively nearby SNRs on the basis of direct observations of proper motions of X-ray-emitting shells. We devised such an 'experiment' by observing the northwest shell of RX J1713.7−3946; this was several years after the initial Chandra observation performed in 2000 (ref. 6).

The new observations with Chandra revealed that the X-ray emission from selected compact regions of the shell seemed variable in flux. The fact that X-ray emission is variable contains unique information about the magnetic field in the particle acceleration region. For a magnetic field typical of the interstellar medium, the acceleration and radiative cooling times of electrons responsible for synchrotron X-ray emission exceed hundreds of years; we therefore 'see' ultrarelativistic electrons, through their nonthermal radiation, accumulated over the history of evolution of the source. The remnant is likely to be about 1,600 years old[15]. The observed rapid variability of X-rays implies that we are dealing with much shorter timescales and provides strong evidence of amplification of the magnetic field around the SNR shell, from the seed (interstellar) value of the order of several μG to the level of mG, as derived below.

We observed the northwest part of RX J1713.7−3946 twice, in July 2005 and May 2006, with the imaging array of the Chandra Advanced CCD Imaging Spectrometer. We also re-analysed the previous Chandra observation of the northwest part in 2000. Figure 1a shows the overall X-ray morphology of the western shell of the SNR, along with sequences of X-ray images of two selected small regions (boxes

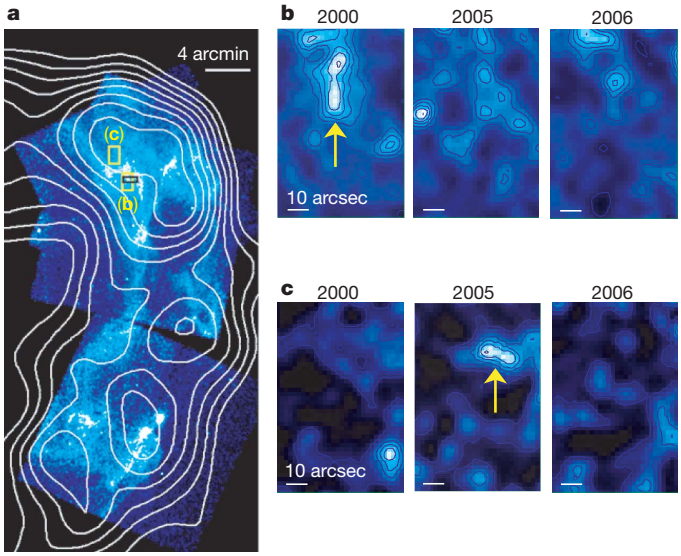


Figure 1 | Chandra X-ray images of the western shell of SNR RX J1713.7−3946. a, A Chandra X-ray mosaic image is overlaid with TeV γ-ray contours from HESS measurements[26]. North is up and east is to the left. With a likely distance to the source $d = 1$ kpc (refs 16, 17), 1 arcmin corresponds to 0.29 pc or 0.95 light years. All Chandra images available are combined into the mosaic image in the energy interval of 1–2.5 keV with a pixel size of 2 arcsec, and smoothed with a gaussian kernel of 8 arcsec to improve the visual appearance. The colour scale is linear in the range $S = (0–1.2) \times 10^{-7}$ photons cm^{-2} s^{-1} per pixel. **b**, A sequence of X-ray observations in July 2000, July 2005 and May 2006 for a small box (labelled b) in **a**. (The black oblong overlapping box b indicates the location in which we compare the cross-sections in Fig. 2.) At each epoch, observations were made for about 30-ks exposure times. The images are produced in the same manner as in **a** and are displayed as surface brightness squared, S^2 (in the range $S = (0–1.6) \times 10^{-7}$ photons cm^{-2} s^{-1} per pixel), to highlight bright regions. A well-defined emission feature with 430 counts ($\sim 10\sigma$ above the surrounding level) in the image from 2000 faded before the year 2005, revealing rapid variability on timescales of years. **c**, Hard-band (3.5–6-keV) images of a box (labelled c) in **a** in terms of S^2 (in the range $S = (0–0.65) \times 10^{-7}$ photons cm^{-2} s^{-1} per pixel). A 'hot spot' (indicated by a yellow arrow) appeared in the 2005 image with 148 counts ($\sim 6\sigma$) and faded before May 2006.

[1]Department of High Energy Astrophysics, Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency (ISAS/JAXA), 3-1-1 Yoshinodai, Sagamihara, Kanagawa, 229-8510, Japan. [2]Dublin Institute for Advanced Studies, 5 Merrion Square, Dublin 2, Ireland. [3]Max-Planck-Institut für Kernphysik, PO Box 103980, D69029 Heidelberg, Germany. [4]Stanford Linear Accelerator Center, 2575 Sand Hill Road, M/S 29, Menlo Park, California 94025, USA.

©2007 Nature Publishing Group

labelled b and c) at the three epochs. It is clear from Fig. 1b that there is a significant decrease in emissivity in a compact region on a timescale of less than a few years. In Fig. 2 we compare the cross-sections of the X-ray images including this decayed feature in 2000, 2005 and 2006.

Another remarkable feature in the images is that a 'hot spot' 20×5 arcsec^2 in size, seen in 2005 with $\sim 6\sigma$ statistical significance, is not present in the 2000 image, as shown in Fig. 1c. This newly emerged feature disappeared again before May 2006, requiring a decay time of less than one year. We found a dozen additional time-varying compact features of lower significance (2–3σ) from various parts of the northwest shell. They are spatially resolved with Chandra, having a typical apparent width of $\sim$4 arcsec.

The original aim of our imaging analysis was to search for proper motions in the shell structure. By comparing the locations of the outer edge of the shell from the first observation in July 2000 and the latest observation in May 2006, we found that the angular displacement between the two epochs does not exceed 6 arcsec, which implies an upper limit on the shock speed v_s of less than 4,500 km s^{-1} for the adopted distance of 1 kpc (refs 16, 17). This estimate is somewhat less restrictive than originally expected, mainly because of the observed variability of X-ray emission.

The time-variability of X-ray flux on a timescale of Δt can, in principle, be revealed only from a compact region smaller than $c\Delta t$ (c being the speed of light); that is, within an angular size $\theta < c\Delta t/d \approx 1$ arcmin for $\Delta t = 1$ year. This is, however, a necessary but not a sufficient condition. The X-ray variability of compact hot spots in the shell requires sufficiently fast production and losses of parent particles. For any set of reasonable parameters characterizing a SNR (such as total energy budget, gas density, radiation and magnetic fields), only synchrotron radiation of ultrarelativistic electrons accelerated at shock fronts can satisfy these conditions. Other emission models discussed in the literature, such as bremsstrahlung of suprathermal electrons or free–free emission of hot thermal plasma, are now safely rejected. Our results therefore provide the strongest observational proof for the synchrotron origin of X-ray emission from RX J1713.7−3946, and in fact for nonthermal X-ray emission detected so far from any SNR.

The decay of the X-ray brightness can be caused by a decrease in total kinetic energy contained in relativistic electrons (W_e), for example as a result of a decrease in the acceleration efficiency while the electrons continue to radiate, if the synchrotron cooling time is of the order of a year or less. The rapid decrease in emission is unlikely to associate with an escape of electrons; whereas the convective escape is too slow, the fast diffusive escape would contradict the very

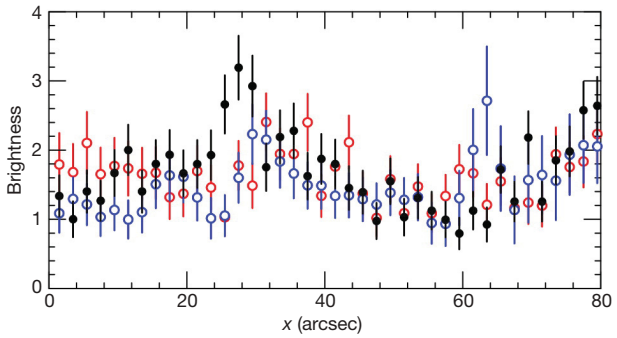


Figure 2 | Cross-sections of the X-ray images at the three epochs. The horizontal (in the east-to-west direction) profiles are extracted from a strip region, indicated as a black oblong overlapping box b in Fig. 1a. Data points are 1–2.5-keV brightness in units of 10^{-6} photons cm^{-2} s^{-1} per bin with a bin size of 2 arcsec with 1σ error bars. They are derived by integrating the surface brightness over a vertical dimension (0.5 arcmin width) for observations in the years 2000 (black), 2005 (red), and 2006 (blue). The brightness peak around $x = 28$ arcsec recorded in 2000 has a clear excess compared with the observations in 2005 and 2006.

requirement of effective acceleration. The radiative cooling time of electrons responsible for a synchrotron photon of energy ε is $t_{\rm synch} \approx 1.5\,(B/{\rm mG})^{-1.5}\,(\varepsilon/{\rm keV})^{-0.5}$ years in a magnetic field of strength B, given that the cooling time of an electron of energy E through synchrotron radiation is $t_{\rm synch} \approx 12\,(B/{\rm mG})^{-2}\,(E/{\rm TeV})^{-1}$ years, and the average energy of synchrotron photons is $\varepsilon \approx 0.016\,(B/{\rm mG})\,(E/{\rm TeV})^2$ keV. Therefore, the average magnetic field in the hot spot should be at the level of more than 1 mG to explain the observed timescale of X-ray disappearance.

The increase in X-ray flux can be caused by an increase in W_e as a result of the boost of the acceleration rate and/or the amplification of the magnetic field. The DSA theory predicts an acceleration time $t_{\rm acc}$ of $\sim 10D(E)v_s^{-2}$ (see, for example, ref. 13), where $D(E) = \eta r_g c/3$ is the diffusion coefficient, $r_g = E/(eB)$ is the gyroradius of an electron with energy E and charge e, and the parameter η characterizes the efficiency of diffusion. The timescale can be rewritten, for electrons emitting synchrotron radiation at energy ε, in the form

$$t_{\rm acc} \approx 1\eta(\varepsilon/{\rm keV})^{0.5}(B/{\rm mG})^{-1.5}(v_s/3{,}000\,{\rm km\,s^{-1}})^{-2}\ {\rm years} \quad (1)$$

For the shock speed $v_s < 4{,}500$ km s^{-1}, as derived in this work, the acceleration time of electrons responsible for X-rays significantly exceeds the timescales of the observed X-ray variability unless the magnetic field upstream is of order 1 mG and particle acceleration proceeds in the maximum effective (Bohm-diffusion) regime with $\eta \approx 1$. Note that equation (1) corresponds to the parallel shock acceleration, the most feasible and generally accepted version of DSA. It has been suggested[18], however, that in strongly oblique shocks the rate of the energy gain of particles could be higher. In addition, it has recently been argued[19] that nonlinear shock-acceleration can proceed beyond the Bohm limit. If so, this would relax somewhat the above requirement to the strength of the magnetic field. In contrast, the estimate of the magnetic field derived from the synchrotron cooling time of electrons does not depend on the electron acceleration mechanism, and therefore the requirement of $B > 1$ mG remains rather robust.

The mG-scale magnetic field inferred from the X-ray variability is evidence in support of the substantial amplification of the magnetic field upstream of the shock from the interstellar value. This is a key condition for the acceleration of protons and nuclei to energies beyond the so-called Lagage–Cesarsky limit around 100 TeV (ref. 20). Recent theoretical studies[2,3] suggest that CR-excited magneto-hydrodynamic waves are indeed able, at least in principle, to amplify the magnetic field by orders of magnitude from its initial seed value, although many complex, highly nonlinear microscopic processes remain to be explored.

The strength of the magnetic field can be estimated indirectly from the width of X-ray filaments if one interprets the origin of these thin structures in terms of diffusion and synchrotron cooling of electrons[21–25]. In particular, for RX J1713.7−3946, this method gives lower limits on the field strength in a range of 0.07–0.25 mG (ref. 25). The variability of the compact hot spots is probably a manifestation of the strongest amplification of the magnetic field, whereas more diffuse regions in the shell would have somewhat weaker magnetic fields.

The broadband X-ray spectroscopy obtained recently by the Suzaku satellite (see Fig. 3) provides independent evidence of very effective acceleration of particles in the shell of RX J1713.7−3946. The measured energy spectrum agrees well with theoretical expectations for the spatially integrated synchrotron spectrum[14] described by a single parameter $\varepsilon_0 = 0.55\eta^{-1}(v_s/3{,}000\,{\rm km\,s^{-1}})^2$ keV. The Suzaku spectral data from 0.4 to 40 keV can be well fitted by the theoretical spectrum with $\varepsilon_0 = 0.55$–0.65 keV, which implies that acceleration should proceed close to the extreme Bohm diffusion limit, $\eta \approx 1$, given the constraint of $v_s < 4{,}500$ km s^{-1}.

Now we can estimate the acceleration time of cosmic-ray protons with energy E at the present epoch merely on the basis of the X-ray

©2007 Nature Publishing Group

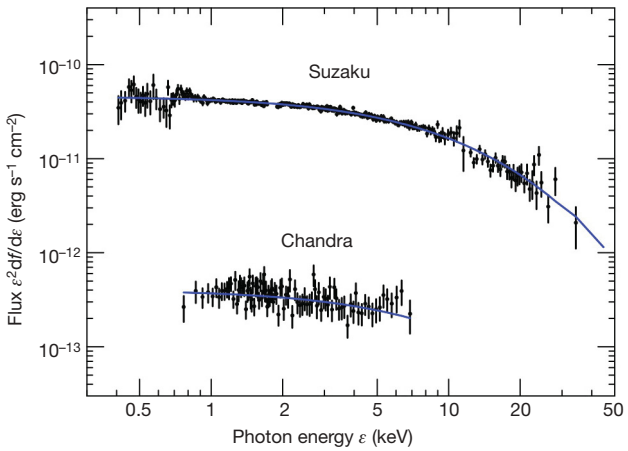


Figure 3 | Energy spectrum of X-ray emission of SNR RX J1713.7−3946. Broadband spectroscopic measurements were made with the Suzaku X-ray satellite in September 2005 (ref. 4). Error bars are 1σ. The Suzaku X-ray spectrum from 0.4 to 40 keV was obtained after integration over the southwest shell of the remnant. In addition, a Chandra X-ray spectrum taken in July 2000 is extracted from the filamentary feature in box b of Fig. 1a. Here we fit the spectra with the theoretical spectrum (solid blue lines) of synchrotron radiation from the shock-accelerated electrons, attenuated by the intervening gas with hydrogen column density $N_{\rm H}$. The synchrotron spectrum predicted by classical DSA theory has the form $\varepsilon^2 {\rm d}f/{\rm d}\varepsilon \propto [1+0.46(\varepsilon/\varepsilon_0)^{0.6}]^{2.29}\exp[-(\varepsilon/\varepsilon_0)^{0.5}]$, with $\varepsilon_0 = 0.55\eta^{-1}(v_s/3{,}000\,{\rm km\,s^{-1}})^2$ keV (ref. 14). The flux points are reconstructed with the best-fit function, with correction for the interstellar absorption. The Suzaku wide-band spectrum has a convex curve, rolling off above ~5 keV. We obtain $\varepsilon_0 = 0.7 \pm 0.2$ keV, with $N_{\rm H} = (0.79 \pm 0.05) \times 10^{22}\,{\rm cm^{-2}}$, for the Chandra spectrum, whereas $\varepsilon_0 = 0.61 \pm 0.05$ keV, with $N_{\rm H} = (0.77 \pm 0.02) \times 10^{22}\,{\rm cm^{-2}}$, for the Suzaku data. The best-fit values of ε_0 yield $\eta \approx 1(v_s/3{,}000\,{\rm km\,s^{-1}})^{-2}$. We find that a good fit to the X-ray spectrum can also be formally obtained with the so-called SRCUT model[10], but only below 10 keV. The limited angular resolution of Suzaku unfortunately does not allow spectral measurements of selected compact regions. However, similar shapes of the Chandra spectra extracted from various regions[6] suggest that the Suzaku spectrum generally also characterizes the broadband X-ray spectra of small-scale regions.

measurements, $t_{\rm acc} \approx 0.1(B/{\rm mG})^{-1}(E/{\rm TeV})$ years, assuming that the acceleration times of protons and electrons with the same energy are equal. Whereas the maximum energy of electrons is limited by synchrotron losses below ~10 TeV, protons and nuclei can be accelerated, over time ΔT, to $E \approx 1(B/{\rm mG})(\Delta T/100\ {\rm years})$ PeV. This is also supported by recent γ-ray observations of RX J1713.7−3946, because the observed spectrum extending to 100 TeV can be naturally explained by hadronic proton–proton interactions[26]. Note that the TeV γ-rays cannot be accounted for by inverse Compton scattering of ultrarelativistic electrons. Indeed, the simple one-zone inverse Compton model, which requires an average magnetic field of ~10 μG (ref. 26), is immediately excluded because of the mG-scale magnetic field derived above. The good spatial correlation of TeV γ-ray and X-ray fluxes as seen in Fig. 1a does not leave room for speculations about a more sophisticated two-zone model. The keV–TeV correlation is in fact better explained by the hadronic model of γ-rays, assuming that protons and electrons are effectively confined in and interacting with the same shocked regions of the SNR shell.

Received 19 June; accepted 30 August 2007.

1. Hillas, A. M. Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays? *J. Phys. G* **31,** R95–R131 (2005).
2. Bell, A. R. & Lucek, S. G. Cosmic ray acceleration to very high energy through the non-linear amplification by cosmic rays of the seed magnetic field. *Mon. Not. R. Astron. Soc.* **321,** 433–438 (2001).
3. Bell, A. R. Turbulent amplification of magnetic field and diffusive shock acceleration of cosmic rays. *Mon. Not. R. Astron. Soc.* **353,** 550–558 (2004).
4. Takahashi, T. *et al.* Measuring the broad-band X-ray spectrum from 400 eV to 40 keV in the southwest part of the supernova remnant RX J1713.7−3946. *Publ. Astron. Soc. Jpn* (in the press); preprint at arXiv:0708.2002v1 [astro-ph].
5. Pfeffermann, E. & Aschenbach, B. in *Röntgenstrahlung from the Universe* (eds Zimmermann, H. U., Trümper, J. E. & Yorke, H.) 267–268 (Report No. 263, Max-Planck-Institut für extraterrestrische Physik, Garching, 1996).
6. Uchiyama, Y., Aharonian, F. A. & Takahashi, T. Fine-structure in the nonthermal X-ray emission of SNR RX J1713.7−3946. *Astron. Astrophys.* **400,** 567–574 (2003).
7. Lazendic, J. S. *et al.* A high-resolution study of nonthermal radio and X-ray emission from supernova remnant G347.3−0.5. *Astrophys. J.* **602,** 271–285 (2004).
8. Cassam-Chenaï, G. *et al.* XMM-Newton observations of the supernova remnant RX J1713.7−3946 and its central source. *Astron. Astrophys.* **427,** 199–216 (2004).
9. Hiraga, J. S., Uchiyama, Y., Takahashi, T. & Aharonian, F. A. Spectral properties of nonthermal X-ray emission from the shell-type SNR RX J1713.7−3946 as revealed by XMM-Newton. *Astron. Astrophys.* **431,** 953–961 (2005).
10. Reynolds, S. P. Models of synchrotron X-rays from shell supernova remnants. *Astrophys. J.* **493,** 375–396 (1998).
11. Drury, L. O'C. *et al.* Test of galactic cosmic-ray source models—working group report. *Space Sci. Rev.* **99,** 329–352 (2001).
12. Blandford, R. & Eichler, D. Particle acceleration at astrophysical shocks: a theory of cosmic ray origin. *Phys. Rep.* **154,** 1–75 (1987).
13. Malkov, M. A. & Drury, L. O'C. Nonlinear theory of diffusive acceleration of particles by shock waves. *Rep. Prog. Phys.* **64,** 429–481 (2001).
14. Zirakashvili, V. N. & Aharonian, F. Analytical solutions for energy spectra of electrons accelerated by nonrelativistic shock-waves in shell type supernova remnants. *Astron. Astrophys.* **456,** 695–702 (2007).
15. Wang, Z. R., Qu, Q.-Y. & Chen, Y. Is RX J1713.7−3946 the remnant of the AD393 guest star? *Astron. Astrophys.* **318,** L59–L61 (1997).
16. Koyama, K. *et al.* Discovery of non-thermal X-rays from the northwest shell of the new SNR RX J1713.7−3946: the second SN 1006? *Publ. Astron. Soc. Jpn.* **49,** L7–L11 (1997).
17. Fukui, Y. *et al.* Discovery of interacting molecular gas toward the TeV gamma-ray peak of the SNR G347.3−0.5. *Publ. Astron. Soc. Jpn.* **55,** L61–L64 (2003).
18. Jokipii, J. R. Rate of energy gain and maximum energy in diffusive shock acceleration. *Astrophys. J.* **313,** 842–846 (1987).
19. Malkov, M. A. & Diamond, P. H. Nonlinear shock acceleration beyond the Bohm limit. *Astrophys. J.* **642,** 244–259 (2006).
20. Lagage, P. O. & Cesarsky, C. J. The maximum energy of cosmic rays accelerated by supernova shocks. *Astron. Astrophys.* **125,** 249–257 (1983).
21. Berezhko, E. G., Ksenofontov, L. T. & Völk, H. J. Confirmation of strong magnetic field amplification and nuclear cosmic ray acceleration in SN 1006. *Astron. Astrophys.* **412,** L11–L14 (2003).
22. Vink, J. & Laming, J. M. On the magnetic fields and particle acceleration in Cassiopeia A. *Astrophys. J.* **584,** 758–769 (2003).
23. Bamba, A. *et al.* Small-scale structure of the SN 1006 shock with Chandra observations. *Astrophys. J.* **589,** 827–837 (2003).
24. Yamazaki, R., Yoshida, T., Terasawa, T., Bamba, A. & Koyama, K. Constraints on the diffusive shock acceleration from the nonthermal X-ray thin shells in SN 1006 NE rim. *Astron. Astrophys.* **416,** 595–602 (2004).
25. Berezhko, E. G. & Völk, H. J. Theory of cosmic ray production in the supernova remnant RX J1713.7−3946. *Astron. Astrophys.* **451,** 981–990 (2006).
26. Aharonian, F. A. *et al.* Primary particle acceleration above 100 TeV in the shell-type supernova remnant RX J1713.7−3946 with deep HESS observations. *Astron. Astrophys.* **464,** 235–243 (2007).

Acknowledgements This work is based on observations made with the Chandra X-ray Observatory, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA. This work also uses data obtained with the Suzaku X-ray satellite, which was developed at the Institute of Space and Astronautical Science of Japan Aerospace Exploration Agency (ISAS/JAXA) in collaboration with NASA, MIT and Japanese institutions.

Author Contributions Y.U. performed analysis of data obtained with the Chandra observations. F.A. investigated theoretical aspects of this work. Y.U. and F.A. jointly wrote the paper. T. Tanaka performed analysis of data obtained with the Suzaku observations. Y.M. checked the analysis of the Chandra data. T. Takahashi wrote a proposal requesting the Suzaku observations. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.U. (uchiyama@astro.isas.jaxa.jp).

©2007 Nature Publishing Group

LETTERS

The a.c. and d.c. Josephson effects in a Bose–Einstein condensate

S. Levy[1], E. Lahoud[1], I. Shomroni[1] & J. Steinhauer[1]

The alternating- and direct-current (a.c. and d.c.) Josephson effects were first discovered in a system of two superconductors, the macroscopic wavefunctions of which are weakly coupled via a tunnelling barrier[1,2]. In the a.c. Josephson effect[1–7], a constant chemical potential difference (voltage) is applied, which causes an oscillating current to flow through the barrier. Because the frequency is proportional to the chemical potential difference only, the a.c. Josephson effect serves as a voltage standard[2]. In the d.c. Josephson effect, a small constant current is applied, resulting in a constant supercurrent flowing through the barrier[4,5,8]. In a sense, the particles do not 'feel' the presence of the tall tunnelling barrier, and flow freely through it with no driving potential. Bose–Einstein condensates should also support Josephson effects[9]; however, while plasma oscillations have been seen[10] in a single Bose–Einstein condensate Josephson junction, the a.c. Josephson effect remains elusive. Here we observe the a.c. and d.c. Josephson effects in a single Bose–Einstein condensate Josephson junction. The d.c. Josephson effect has been observed previously only in superconducting systems[11]; in our study, it is evident when we measure the chemical potential–current relation of the Bose–Einstein condensate Josephson junction[4,11]. Our system constitutes a trapped-atom interferometer[12,13] with continuous readout[14], which operates on the basis of the a.c. Josephson effect. In addition, the measured chemical potential–current relation shows that the device is suitable for use as an analogue of the superconducting quantum interference device, which would sense rotation[2,15].

We can understand all of the effects reported here from the following three nonlinear equations describing the model of two Bose–Einstein condensates (BECs)[7], which are weakly coupled in a double-well potential, in the presence of dissipation due to thermal atoms (see Supplementary Information). These equations are derived from the Gross–Pitaevskii equation (GPE)[16], with the exception of the second term of equation (1), which represents the effect of non-zero temperature. The atomic current $\dot{\eta}$ flowing across the barrier is given by[17]:

$$\dot{\eta} = \omega_J \sin\phi - G\Delta\mu \quad (1)$$

where $\eta \equiv (N_1 - N_2)/N$ is the relative population between the two condensates with $N \equiv N_1 + N_2$, $\phi \equiv \phi_1 - \phi_2$ is the relative phase between the two condensates, $N\hbar\omega_J/2$ is the Josephson coupling energy, and G is the conductance due to thermal atoms[5,17,18]. The chemical potential difference $\Delta\mu \equiv \mu_1 - \mu_2$ is given by[4]:

$$\Delta\mu = \hbar\omega_C\left(\eta - \eta_{equil}\right) \quad (2)$$

where $2\hbar\omega_C/N$ is the capacitive energy, and η_{equil} is the equilibrium value of η, which is non-zero for a double well with different volumes on the two sides. Equation (2) expresses the fact that a number imbalance is associated with a chemical potential difference, due to repulsive interactions. The phase evolution is given by[1]:

$$\dot{\phi} = -\frac{\Delta\mu}{\hbar} \quad (3)$$

Equation (3) is valid for any two well-separated condensates ($\omega_J \ll \omega_C$), whether they are coupled by tunnelling or not[4,19]. This phase evolution has been observed[12,13] in the Fock regime of completely uncoupled condensates ($\omega_C/\omega_J \gg (N/2)^2$), as well as for condensates in the Josephson regime ($1 \ll \omega_C/\omega_J \ll (N/2)^2$), for the case of $\omega_C/\omega_J \approx 15$ (ref. 10). The Josephson regime characterizes the current experiment, as well as the superconducting case[19]. Equations (1)–(3) describe each of the three nonlinear systems shown in Fig. 1 (ref. 2). These systems are analogous to the BEC Josephson junction, and will provide an intuitive description of the effects observed.

We will now discuss the various solutions of equations (1)–(3), beginning with the a.c. Josephson effect. Consider equations (1) and (3), with a constant $\Delta\mu$ applied. Then $\dot{\phi}$ is constant, and for $G = 0$, the current becomes $\dot{\eta} = -\omega_J \sin\left(\frac{\Delta\mu}{\hbar}t\right)$ (refs 2, 3). Thus, the current oscillates in time with frequency $\Delta\mu/\hbar$, as indicated by the solid line of Fig. 2a. This a.c. Josephson effect relies on the sinusoidal nature of equation (1), which results from the weak coupling between the two macroscopic wavefunctions[20]. In contrast, in a bulk macroscopic quantum system, the current is proportional to the gradient of the phase, so equation (1) would be replaced by $\dot{\eta} \propto \phi$. Such a system would show no oscillating current because ϕ increases at a constant rate. As shown in Fig. 1a, the a.c. Josephson effect corresponds to the rigid pendulum executing complete rotations, with an angle that always increases[2]. The period of rotation of the pendulum corresponds to the period of the current oscillations, $(\Delta\mu/h)^{-1}$.

The a.c. Josephson effect is very different from the plasma oscillations indicated in Fig. 1a, for which the rigid pendulum oscillates resonantly about zero with $-\pi < \phi < \pi$ (ref. 21). The oscillations observed in ref. 10 were plasma oscillations, as were the oscillations in the lattice of ref. 22. These plasma oscillations exist even for small-angle oscillations, for which equation (1) reduces to $\dot{\eta} \propto \phi$, as in a bulk macroscopic quantum system. The a.c. Josephson effect, in contrast, is a striking demonstration of the sinusoidal nature of equation (1), and has no analogue in a bulk system.

The a.c. Josephson effect has been seen in superfluid ^{3}He (ref. 23), as well as in superconductors. In BEC, this effect has been seen in systems where the chemical potential difference $\Delta\mu$ is produced by an asymmetric potential rather than by a population imbalance. This includes a lattice of Josephson junctions[24], as well as an internal a.c. Josephson effect[25]. In the single BEC Josephson junction reported here, $\Delta\mu$ is produced by a population imbalance, as in the superconducting case.

Our ultrahigh-resolution BEC apparatus applies optical potentials through an aspheric lens with a numerical aperture of 0.5, as shown in Fig. 3a. The imaging is performed through this same lens. The

[1]Department of Physics, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel.

©2007 Nature Publishing Group

observed resolution for creating potentials as well as imaging is 1.2 μm. In the current experiment, the apparatus is used to create a double-well Josephson junction, as shown in Fig. 3b. The tunnelling barrier of the double well is a highly elongated laser beam, with a $1/e^2$ diameter of 1.4 μm along the tunnelling axis and 400 μm along the long axis. The laser's wavelength of 774 nm is blue-detuned from resonance (780 nm). The barrier divides the axially symmetric, harmonic magnetic trap (measured trap frequencies of $\omega_r/2\pi = 224$ Hz and $\omega_z/2\pi = 26$ Hz) parallel to its long axis. The position of the centre of the harmonic trap relative to the barrier is controlled by an additional magnetic coil. After transporting the ^{87}Rb atoms in the ($F = 2$, $m_F = 2$) state to the high-resolution region via a moving quadrupole trap[26], evaporation is performed in the double well, resulting in $N = 10^5$ atoms in the two condensates.

To observe the a.c. Josephson effect, we create a chemical potential difference $\Delta\mu$ between the two clouds, and measure the frequency of the oscillations which accompany the almost constant $\Delta\mu$. First, BEC is reached by evaporating into an asymmetric double well potential, as shown in the inset of Fig. 2a. The harmonic trap is then shifted over 5 to 10 ms to form a symmetric double well, but with a chemical potential difference, as shown in the inset. After the shift, twelve non-destructive phase contrast images are taken of the BEC Josephson junction, giving $\Delta\mu$ and $\eta(t)$, as shown in Fig. 4a–c. The oscillation frequencies of the two curves of Fig. 4c are seen to be proportional to $\Delta\mu$. To find the frequency of oscillation ω, the Fourier transform of $\eta(t)$ is computed, and the peak of the Fourier transform is taken as ω (ref. 23). Figure 2a shows the resulting measurement of the a.c. Josephson effect in a BEC. The measured frequency $\dot{\phi}$ of the atomic current gives the chemical potential difference via equation (3).

Although the a.c. Josephson effect has not been observed previously in a BEC Josephson junction, a related effect has been seen, which we now describe. In the absence of dissipation ($G = 0$), the speed of the freely rotating rigid pendulum averaged over a complete rotation never decreases, which implies that the population imbalance in the a.c. Josephson effect, averaged over a complete period, never decreases[19], as indicated by the upper, blue solid curve of Fig. 2b. In other words, in the absence of dissipation, no net current flows in the a.c. Josephson effect, only an alternating current. This non-equilibration associated with the underlying a.c. Josephson effect is referred to as "macroscopic quantum self-trapping" (MQST)[7,19], which was observed in a BEC[10]. In Fig. 1b, the freely rotating rigid pendulum is equivalent to $\dot{\eta}_{equil} = 0$, with a charged capacitor. MQST reflects the fact that owing to the a.c. Josephson effect, the capacitor will never discharge.

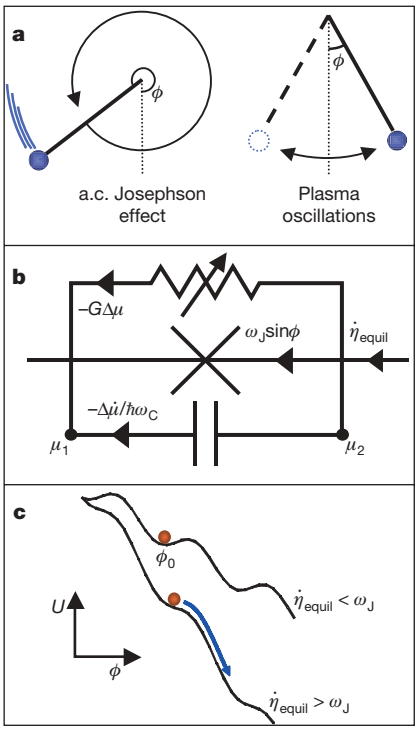


Figure 1 | Nonlinear systems analogous to the BEC Josephson junction. Each of these systems represents equations (1)–(3). The damping time of each of these systems is of the order of $(\hbar G\omega_c)^{-1}$. The conductance G is varied by adjusting the fraction of thermal atoms. **a**, The rigid pendulum. The case of zero current bias ($\dot{\eta}_{equil} = 0$) is shown. The relative phase ϕ is represented by the angle of the rigid pendulum, and the population imbalance η is proportional to the speed $\dot{\phi}$. **b**, Superconducting circuit, analogous to the resistively shunted Josephson junction model. The ×-shape represents the Josephson junction. The current through each of the branches is indicated on the figure. The constant applied current bias is given by $\dot{\eta}_{equil}$. **c**, A particle moving in a washboard potential. By computing the energy $\int \Delta\mu\, dN$ for each of the currents in **b**, we find that the system is equivalent to a particle of mass $N\hbar/2\omega_C$ moving in a washboard potential given by $U = (\hbar N/2)\left[\omega_J(1-\cos\phi) - \dot{\eta}_{equil}\phi\right]$. The tilt of the washboard potential is due to the current bias $\dot{\eta}_{equil}$.

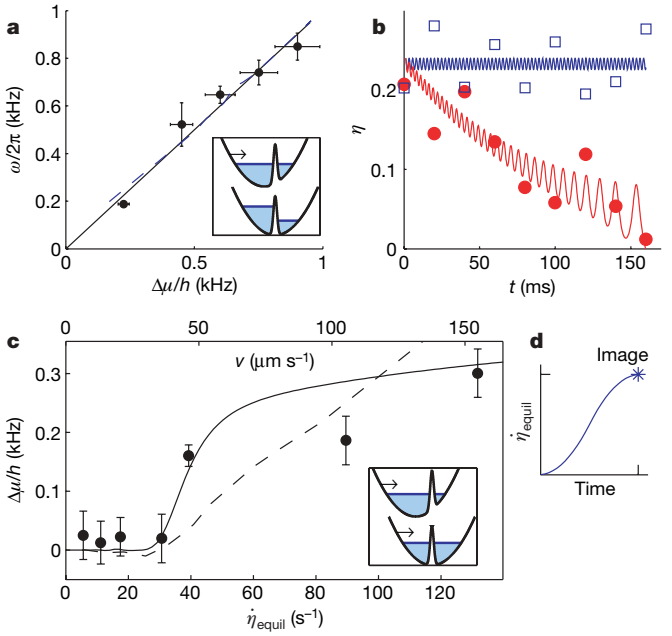


Figure 2 | Observation of the a.c. and d.c. Josephson effects. **a**, The a.c. Josephson effect. The solid line shows $\omega/2\pi = \Delta\mu/h$, with no free parameters. The dashed curve is a full simulation of the GPE. Error bars show the standard error of the mean, for between two and twelve repetitions of the experiment. The inset shows the creation of $\Delta\mu$. **b**, The decay of the MQST. The measured blue squares show a long damping time, with a thermal fraction of roughly 5% ($T \approx 0.3T_c$) (ref. 16). The increased damping seen for the red circles is with a thermal fraction of roughly 20% ($T \approx 0.5T_c$), obtained by increasing the final evaporation frequency by 2.5 kHz. The solid curves are numerical solutions of equations (1)–(3). In the decaying red curve, $(\hbar G\omega_c)^{-1}$ is adjusted to 0.1 s. **c**, The $\Delta\mu$–I relation and the d.c. Josephson effect. Filled circles are the measured values, which show the d.c. Josephson effect for $\dot{\eta}_{equil} < 30\ s^{-1}$. Three of the points are averages of two measurements. For these points the error bars indicate the standard error of the mean. The average of these error bars, multiplied by $\sqrt{2}$, gives the error bars shown for the other points. The solid line is a numerical solution of equations (1)–(3). The dashed curve is a GPE simulation, in which the barrier height is adjusted to a value of (2.5 kHz)×h. The inset shows the application of $\dot{\eta}_{equil}$. **d**, Before imaging each point in **c**, $\dot{\eta}_{equil}$ is increased slowly, to prevent plasma oscillations in the potential of Fig. 1c. Any plasma oscillations around ϕ_o would allow the system to leave the supercurrent state, and enter the finite $\Delta\mu$ state, even for $\dot{\eta}_{equil} < \omega_J$. With $\dot{\eta}_{equil}$ at its final value, which occurs at the moment that the potential well is symmetric, we image the atoms and determine $\Delta\mu$ by the same technique shown in Fig. 4.

©2007 Nature Publishing Group

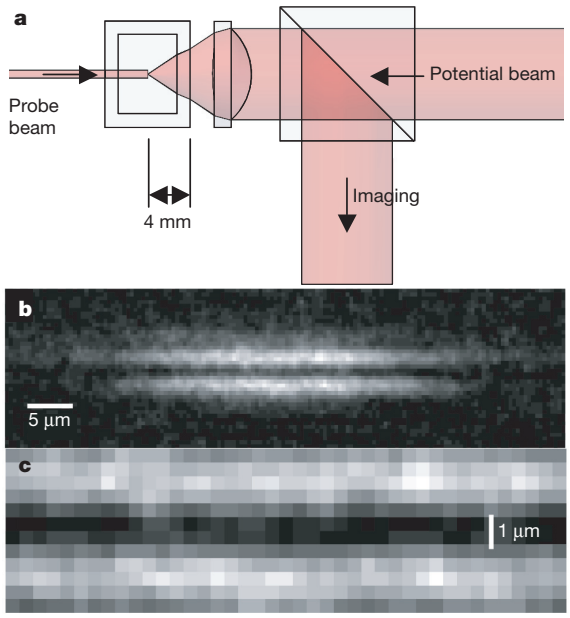


Figure 3 | Creating and imaging our BEC Josephson junction. a, The application of high-resolution potentials, and imaging with the same resolution. Almost arbitrary potentials can be formed by engineering the potential beam. The left rectangular region is the cross-section of the vacuum cell. **b**, *In situ* image of our BEC Josephson junction. The light regions are the two BECs. The image is obtained with destructive phase contrast imaging, employing small detuning to improve the signal to noise ratio. **c**, An enlargement of the narrow tunnelling region of the wavefunction. Each pixel is 0.46 μm. Gravity is directed downwards in all panels.

However, we observe that if damping due to the thermal atoms is non-zero ($G>0$), the speed of the rigid pendulum will slowly decrease, corresponding to decay of the MQST[5,18,27], as indicated by the lower, red solid curve of Fig. 2b. This decay occurs on a timescale determined by the fraction of thermal atoms, as indicated by the two sets of measured points in Fig. 2b. This timescale is much longer than the time between images in a measurement such as Fig. 4a, which varies between 0.3 and 1.7 ms. The decay of the MQST shows that the thermal cloud significantly modifies the behaviour of the two condensates. This can be understood from equations (1)–(3). The second term of equation (1) implies that $\Delta\mu$, which is mostly due to the condensate atoms, drives the motion of the thermal atoms[5,17,18]. Furthermore, by equations (2) and (3), the resulting population change in the thermal atoms contributes to $\Delta\mu$ and drives the phase evolution of the condensate.

We now turn to the d.c. Josephson effect[1,4,5,8,11], which can be understood in terms of the washboard potential shown in Fig. 1c. As long as the potential is not tilted too much by the applied current $\dot{\eta}_{\mathrm{equil}}$, the system can sit in a minimum with phase ϕ_{o} given by $\dot{\eta}_{\mathrm{equil}}=\omega_{\mathrm{J}}\sin\phi_{\mathrm{o}}$. Because $\dot{\phi}$ is zero, $\Delta\mu$ is zero by equation (3). Thus, a constant supercurrent $\dot{\eta}=\dot{\eta}_{\mathrm{equil}}$ flows through the barrier with no driving potential difference[4,11,21] (the d.c. Josephson effect), as indicated by the solid black curve of Fig. 2c for small $\dot{\eta}_{\mathrm{equil}}$. On the other hand, when $\dot{\eta}_{\mathrm{equil}}$ exceeds the critical current ω_{J}, some current must flow through the resistive or capacitive paths shown in Fig. 1b. These currents result in a non-zero $\Delta\mu$. The relation between the applied current bias $\dot{\eta}_{\mathrm{equil}}$ and the resulting $\Delta\mu$ (the $\Delta\mu$–I relation) thus has both a supercurrent state and a finite $\Delta\mu$ state, with a transition between the states at $\dot{\eta}_{\mathrm{equil}}=\omega_{\mathrm{J}}$, as indicated by the solid curve of Fig. 2c. The two states are also represented by the two curves of Fig. 1c.

To measure the $\Delta\mu$–I relation, we essentially perform the experiment suggested in ref. 4. A bias current $\dot{\eta}_{\mathrm{equil}}$ is applied by moving the harmonic potential at a speed v relative to the barrier, by means of the additional magnetic coil, as shown in the inset of Fig. 2c. The total distance moved is 0.7 μm. Figure 2c shows the resulting measurement of the $\Delta\mu$–I relation of the BEC Josephson junction.

The supercurrent state is seen in Fig. 2c for $\dot{\eta}_{\mathrm{equil}}<30\,\mathrm{s}^{-1}$ ($v<40\,\mu\mathrm{m\,s}^{-1}$). This state represents the d.c. Josephson effect[4,5,8]. The thus measured critical current of $30\,\mathrm{s}^{-1}$ is ω_{J} from equation (1). By comparing $\Delta\mu$ and η, and using equation (2), we find a value

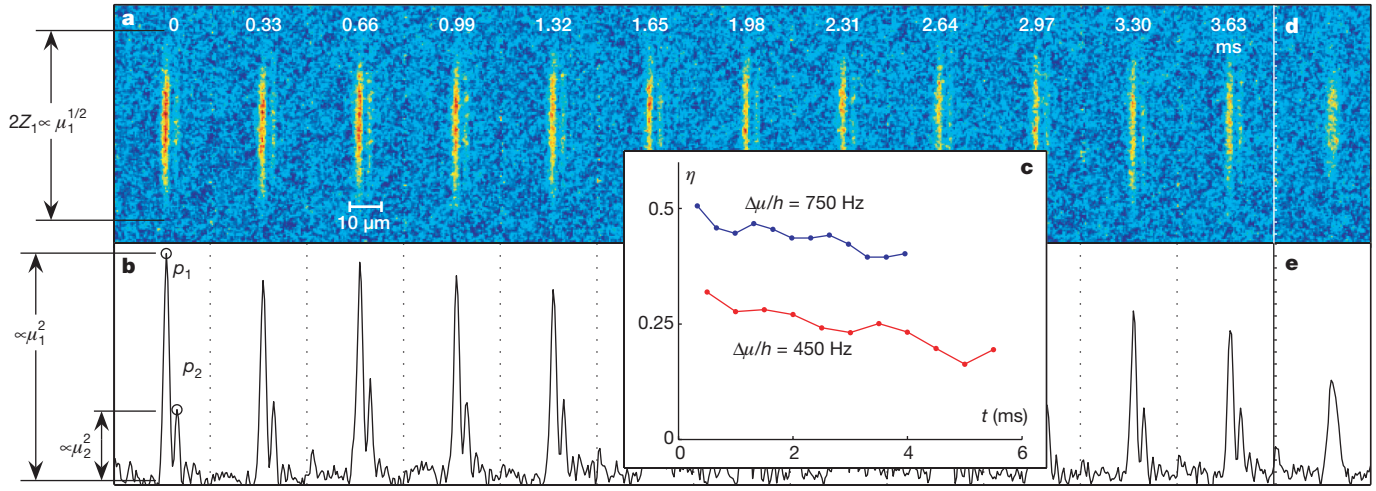


Figure 4 | Time evolution of a BEC Josephson junction. a, Twelve *in situ* images of the same BEC Josephson junction, obtained by non-destructive phase contrast imaging. To obtain an absolute, quantitative value μ_1 for the chemical potential of one of the clouds, its length $2Z_1$ is measured directly from a fit to the image, as shown. This gives the chemical potential from the relation $\mu_1=m\omega_z^2Z_1^2/2$, where m is the atomic mass. The number of atoms N_1 or N_2 is computed by integrating over each cloud, giving $\eta(t)$ for each of the twelve images. **b**, The image integrated in the z-direction. The height of the peaks p_1 and p_2 are proportional to μ_1^2 and μ_2^2, respectively, in the Thomas–Fermi approximation[16]. The chemical potential difference is thus given by $\Delta\mu=\mu_1-\mu_2$ where $\mu_2=\mu_1\sqrt{p_2/p_1}$. **c**, The phase evolution of the BEC Josephson junction. The top, blue curve shows $\eta(t)$, when averaged with seven other repetitions of the experiment. For comparison, the lower, red curve shows a smaller oscillation frequency corresponding to a smaller value of $\Delta\mu$. The slope of the curves is due to imaging losses. Panels **d** and **e** show the BEC in the harmonic trap. After the images shown in **a** and **b** are obtained, the barrier is slowly decreased to zero. The image of the resulting BEC shows the profile of the harmonic trap, via the Thomas–Fermi approximation[16]. We use this profile to verify the alignment of the centre of the harmonic trap with the barrier.

©2007 Nature Publishing Group

of $\omega_C = 9{,}000\,s^{-1}$. Because $\omega_C/\omega_J = 300$, the system is in the Josephson regime: $1 \ll \omega_C/\omega_J \ll (N/2)^2$. As the current increases beyond the critical current in Fig. 2c, the system enters the finite $\Delta\mu$ state. This transition is expected to occur via the disappearance of the minimum in Fig. 1c rather than via fluctuations, because the height of the effective potential barrier of Fig. 1c is on the order of $\hbar N\omega_J/k_B \approx 20\,\mu K$. This is orders of magnitude larger than the actual temperature ($<100\,nK$), or the effective temperature of quantum fluctuations, which is of the order of the plasma oscillation frequency in the washboard potential $\hbar\sqrt{\omega_C\omega_J}/k_B \approx 4\,nK$ (refs 21, 28).

Our BEC Josephson junction can be considered to be a trapped-atom interferometer[12–14]. Like the system of ref. 14, our system gives a continuous measurement of the relative phase between the two coherent clouds of atoms, as shown in Fig. 4c. However, the operation of our interferometer is based on the a.c. Josephson effect, rather than on the non-tunnelling technique of ref. 14. We verify the functioning of the interferometer by measuring various interaction energies, as shown in Fig. 2a. Our interferometer could also be used to measure the time evolution of applied optical or magnetic potentials[12], or to measure a time-varying scattering length. If our interferometer were to be used for a relatively long measurement, it would be essential that $\Delta\mu$ did not equilibrate through the tunnelling barrier. Fortunately, this occurs naturally, owing to the MQST, which occurs for sufficiently low temperatures.

Furthermore, two BEC Josephson junctions could be combined in a loop, to form the analogue of a superconducting quantum interference device (SQUID), which would sense rotation. Such a BEC SQUID would require a macroscopic wavefunction around the entire loop, because it relies on the quantization of phase around the loop. The BEC SQUID would have a $\Delta\mu$–I relation as in Fig. 2c, except that the critical current $\dot{\eta}_c$ would vary with the rate of rotation. The sensitivity to rotation would be proportional to $\delta\dot{\eta}_c/\dot{\eta}_c$, where $\delta\dot{\eta}_c$ is the smallest measurable change in $\dot{\eta}_c$. It is interesting to note that the $\Delta\mu$–I relation of Fig. 2c shows a rapid variation in $\Delta\mu$ for a small variation in $\dot{\eta}_{equil}$. Thus, by measuring $\Delta\mu$ in the BEC SQUID, one could measure the change in $\dot{\eta}_c$, and rotation, with high sensitivity[2,15]. This is analogous to a superconducting SQUID's sensitivity to magnetic flux[2]. A SQUID detector of rotation was achieved in superfluid ^{3}He (ref. 29), but the detection did not exploit the sensitivity gain of the $\Delta\mu$–I relation. A BEC SQUID differs from a Sagnac-effect atom interferometer in that it requires a macroscopic wavefunction around the loop, and $2\delta\dot{\eta}_c/\dot{\eta}_c$ replaces the minimum measurable phase[30]. In principle, owing to the $\Delta\mu$–I relation, a BEC SQUID could have more sensitivity than a Sagnac atom interferometer of the same enclosed area.

Our BEC Josephson junction is the first application of our ultra-high-resolution BEC system, allowing for almost arbitrary potentials and imaging on a tunnelling length scale. We report the measurements of the a.c. and d.c. Josephson effects in a BEC, as well as the measurement of the $\Delta\mu$–I relation. Our BEC Josephson junction can be considered to be an atom interferometer, in which the MQST has the important role of maintaining the chemical potential difference over time. The thermal cloud is seen to significantly alter the dynamics, in that it can cause the decay of the MQST. Furthermore, this system is seen to be suitable for forming the analogue of a SQUID detector, which would sense rotation. The work reported here represents the start of a new type of cold atom experimentation, where the tunnelling can be observed *in situ*, without opening the confining potential.

METHODS SUMMARY

The aspheric lens of Fig. 3a is of a type typically used for coupling a laser beam to a fibre. Specifically, it is a Thorlabs 350240-B, with a focal length of 8 mm. The laser for the potential beam is a Ti:sapphire ring laser (Coherent 899-21), pumped by a frequency-doubled Nd:YVO_4 laser (Coherent Verdi V-10). The long axis of the harmonic trap is oriented perpendicularly to the image of Fig. 3a.

In Fig. 4a, twelve images were obtained of the same BEC Josephson junction, where each image gives the number of atoms in each of the two weakly coupled BECs. Making these measurements of the number of atoms, with a precision limited by shot noise, does not randomize the relative phase. For the large number of atoms involved, the relative number and phase can be considered simultaneously well defined[17,19].

Creating Fig. 4b involved integrating the entire image into a small number of points, which greatly reduces noise. Therefore, we used Fig. 4b to find the ratio between the chemical potentials of the two clouds. The relation $p_{1,2} \propto \mu_{1,2}^2$ in Fig. 4b stems from the parabolic nature of the trap in two dimensions. To obtain an absolute calibration, however, we use the length of the cloud in Fig. 4a. We use the left cloud because it is longer and has more atoms, so we can measure its length more accurately. We find that this method minimizes the noise, while maintaining an absolute calibration of $\Delta\mu$.

To compute the lower horizontal axis from the upper horizontal axis in Fig. 2c, we need the ratio between $\dot{\eta}_{equil}$ and v, which is the ratio between η_{equil} and x. We compute this ratio $d\eta_{equil}/dx$ by computing the ground-state wavefunction by the GPE for a given x. We find $d\eta_{equil}/dx = 0.85\,\mu m^{-1}$. In Fig. 2d, $\dot{\eta}_{equil}$ is ramped up in a time varying between 8.5 and 200 ms. The total distance moved is always 0.7 μm, so $\dot{\eta}_{equil}$ and v are determined by the time of the ramp.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 May; accepted 15 August 2007.

1. Josephson, B. D. Possible new effects in superconductive tunnelling. *Phys. Lett.* **1,** 251–253 (1962).
2. Barone, A. & Paterno, G. *Physics and Applications of the Josephson Effect* Chs 1, 6, 11, 13 (Wiley, New York, 1982).
3. Dalfovo, F., Pitaevskii, L. & Stringari, S. Order parameter at the boundary of a trapped Bose gas. *Phys. Rev. A.* **54,** 4213–4217 (1996).
4. Giovanazzi, S., Smerzi, A. & Fantoni, S. Josephson effects in dilute Bose-Einstein condensates. *Phys. Rev. Lett.* **84,** 4521–4524 (2000).
5. Meier, F. & Zwerger, W. Josephson tunneling between weakly interacting Bose-Einstein condensates. *Phys. Rev. A.* **64,** 033610 (2001).
6. Raghavan, S., Smerzi, A., Fantoni, S. & Shenoy, S. R. Coherent oscillations between two weakly-coupled Bose-Einstein condensates: Josephson effects, π oscillations, and macroscopic quantum self-trapping. *Phys. Rev. A.* **59,** 620–633 (1999).
7. Smerzi, A., Fantoni, S., Giovanazzi, S. & Shenoy, S. R. Quantum coherent atomic tunneling between two trapped Bose-Einstein condensates. *Phys. Rev. Lett.* **79,** 4950–4953 (1997).
8. Sakellari, E., Leadbeater, M., Kylstra, N. J. & Adams, C. S. Josephson spectroscopy of a dilute Bose-Einstein condensate in a double-well potential. *Phys. Rev. A* **66,** 033612 (2002).
9. Javanainen, J. Oscillatory exchange of atoms between traps containing Bose condensates. *Phys. Rev. Lett.* **57,** 3164–3166 (1986).
10. Albiez, M. *et al.* Direct observation of tunneling and nonlinear self-trapping in a single Bosonic Josephson junction. *Phys. Rev. Lett.* **95,** 010402 (2005).
11. Anderson, P. W. & Rowell, J. M. Probable observation of the Josephson superconducting tunneling effect. *Phys. Rev. Lett.* **10,** 230–232 (1963).
12. Shin, Y. *et al.* Atom interferometry with Bose-Einstein condensates in a double-well potential. *Phys. Rev. Lett.* **92,** 050405 (2001).
13. Schumm, T. *et al.* Matter-wave interferometry in a double well on an atom chip. *Nature Phys.* **1,** 57–62 (2005).
14. Saba, M. *et al.* Light scattering to determine the relative phase of two Bose-Einstein condensates. *Science* **307,** 1945–1948 (2005).
15. Packard, R. E. & Vitale, S. Principles of superfluid-helium gyroscopes. *Phys. Rev. B* **46,** 3540–3549 (1992).
16. Dalfovo, F., Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of Bose-Einstein condensation in trapped gases. *Rev. Mod. Phys.* **71,** 463–512 (1999).
17. Zapata, I., Sols, F. & Leggett, A. J. Josephson effect between trapped Bose-Einstein condensates. *Phys. Rev. A.* **57,** R28–R31 (1998).
18. Marino, I., Raghavan, S., Fantoni, S., Shenoy, S. R. & Smerzi, A. Bose-condensate tunneling dynamics: momentum-shortened pendulum with damping. *Phys. Rev. A.* **60,** 487–493 (1999).
19. Leggett, A. J. Bose-Einstein condensation in the alkali gases: some fundamental concepts. *Rev. Mod. Phys.* **73,** 307–356 (2001).
20. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. 3, Ch. 21 (Addison-Wesley, Reading, Massachusetts, 1965).
21. Likharev, K. K. *Dynamics of Josephson Junctions and Circuits* Chs 1, 3, 5 (Gordon and Breach Science Publishers, New York, 1986).
22. Cataliotti, F. S. *et al.* Josephson junction arrays with Bose-Einstein condensates. *Science* **293,** 843–846 (2001).
23. Pereverzev, S. V., Loshak, A., Backhaus, S., Davis, J. C. & Packard, R. E. Quantum oscillations between two weakly coupled reservoirs of superfluid ^{3}He. *Nature* **388,** 449–451 (1997).
24. Anderson, B. P. & Kasevich, M. A. Macroscopic quantum interference from atomic tunnel arrays. *Science* **282,** 1686–1689 (1998).

©2007 Nature Publishing Group

25. Hall, D. S., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Measurements of relative phase in two-component Bose-Einstein condensates. *Phys. Rev. Lett.* **81,** 1543–1546 (1998).
26. Greiner, M., Bloch, I., Hänsch, T. W. & Esslinger, T. Magnetic transport of trapped cold atoms over a large distance. *Phys. Rev. A* **63,** 031401(R) (2001).
27. Ruostekoski, J. & Walls, D. F. Bose-Einstein condensate in a double-well potential as an open quantum system. *Phys. Rev. A* **58,** R50–R53 (1998).
28. Pitaevskii, L. & Stringari, S. Thermal versus quantum decoherence in double well trapped Bose-Einstein condensates. *Phys. Rev. Lett.* **87,** 180402 (2001).
29. Simmonds, R. W., Marchenkov, A., Hoskinson, E., Davis, J. C. & Packard, R. E. Quantum interference of superfluid ^{3}He. *Nature* **412,** 55–58 (2001).
30. Gustavson, T. L., Bouyer, P. & Kasevich, M. A. Precision rotation measurements with an atom interferometer gyroscope. *Phys. Rev. Lett.* **78,** 2046–2049 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank W. Ketterle, M. Segev, E. Altman, Y. Kafri, E. Akkermans, E. Polturak, B. Shapiro and R. Pugatch for readings of the manuscript. This work was supported by the Israel Science Foundation and the Russell Berrie Nanotechnology Institute.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.S. (jeffs@physics.technion.ac.il).

©2007 Nature Publishing Group

METHODS

The harmonic magnetic trap in the high-resolution region is a QUIC-trap[31], which is the coil configuration typically used in conjunction with a moving quadrupole trap[26]. The quadrupole coils of the QUIC-trap are located above and below the vacuum cell in Fig. 3a. The vacuum cell is anti-reflection coated on the outside only.

In ref. 14, a trapped-atom interferometer with continuous measurement was achieved by a non-tunnelling technique. There, the measurement of the phase only perturbed the interferometer through the extraction of a very small sample of atoms. Here, the measurement only perturbs the interferometer through the slight scattering associated with phase-contrast imaging[32].

The imaging process integrates the density along the y direction (perpendicular to the image plane), so Fig. 4b is integrated in both the y and z directions, for both of which the trap is parabolic. The Thomas–Fermi radii in both of these directions are proportional to $\sqrt{\mu_1}$ and $\sqrt{\mu_2}$ for clouds 1 and 2, respectively. This implies that the height of the peaks p_1 and p_2 are proportional to μ_1^2 and μ_2^2, respectively. This is verified by a GPE simulation.

In the measurement of the $\Delta\mu$–I relation, the position of the harmonic trap relative to the barrier is adjusted by applying a current I_e to an additional magnetic coil. To calibrate the upper horizontal axis of Fig. 2c, we apply a known current I_e through the coil, and observe the distance x over which the cloud shifts in the image. For a given current ramp $I_e(t)$, we thus know $v(t)$ of the upper axis, because $v(t) = \frac{dx}{dI_e}\frac{dI_e(t)}{dt}$.

31. Esslinger, T., Bloch, I. & Hänsch, T. W. Bose-Einstein condensation in a quadrupole-Ioffe configuration trap. *Phys. Rev. A.* **58,** R2664–R2667 (1998).
32. Andrews, M. R. *et al.* Direct, nondestructive observation of a Bose condensate. *Science* **273,** 84–87 (1996).

©2007 Nature Publishing Group

Vol 449|4 October 2007|doi:10.1038/nature06182

LETTERS

Fluctuating superconductivity in organic molecular metals close to the Mott transition

Moon-Sun Nam[1], Arzhang Ardavan[1], Stephen J. Blundell[1] & John A. Schlueter[2]

On cooling through the transition temperature T_c of a conventional superconductor, an energy gap develops as the normal-state charge carriers form Cooper pairs; these pairs form a phase-coherent condensate that exhibits the well-known signatures of superconductivity: zero resistivity and the expulsion of magnetic flux (the Meissner effect[1]). However, in many unconventional superconductors, the formation of the energy gap is not coincident with the formation of the phase-coherent superfluid. Instead, at temperatures above the critical temperature a range of unusual properties, collectively known as 'pseudogap phenomena', are observed[2]. Here we argue that a key pseudogap phenomenon—fluctuating superconductivity occurring substantially above the transition temperature—could be induced by the proximity of a Mott-insulating state. The Mott-insulating state in the κ-(BEDT-TTF)$_2$X organic molecular metals[3–5] can be tuned, without doping, through superconductivity into a normal metallic state as a function of the parameter t/U, where t is the tight-binding transfer integral characterizing the metallic bandwidth and U is the on-site Coulomb repulsion. By exploiting a particularly sensitive probe of superconducting fluctuations, the vortex-Nernst effect, we find that a fluctuating regime develops as t/U decreases and the role of Coulomb correlations increases.

So far, pseudogap phenomena have been studied almost exclusively within one class of correlated superconductors, the high-T_c copper oxides, but the interpretation of observations in these materials is hindered by several factors: a very high magnetic field is needed to reach the normal state, the high temperature scale introduces thermal effects which may obscure the intrinsic physics, and the key parameter for tuning the superconducting properties, hole-doping, is inevitably associated with the introduction of chemical and structural disorder. A broader understanding of the underlying physics of the pseudogap would be achieved by identifying the phenomenological similarities across different classes of materials. This motivates our study of the κ-(BEDT-TTF)$_2$X organic molecular metals, which exhibit smaller temperature and magnetic field scales than the copper oxides, a phase diagram that is traversed without the introduction of disorder, and some existing evidence of pseudogap-like behaviour[6,7].

Figure 1 shows the crystal structure and the phase diagram of the κ-(BEDT-TTF)$_2$X family of isostructural organic molecular metals. The flat BEDT-TTF molecules dimerize to form molecular units that stack in planes on a triangular lattice[3]. The anions, X, separate the planes and accept one electron from each BEDT-TTF dimer. The resulting quasi-two-dimensional BEDT-TTF tight-binding band is half-filled. As the anion is modified from X = Cu[N(CN)$_2$]Cl, through X = Cu[N(CN)$_2$]Br to X = Cu(NCS)$_2$, the separation of BEDT-TTF dimers decreases, increasing the tight-binding transfer integral t and increasing the resulting bandwidth[8]; a very similar result may be achieved by applying hydrostatic pressure[5,9]. The on-site Coulomb repulsion U is a property of the dimer, and is almost independent of anion or pressure.

For small t/U, where Coulomb correlations dominate the ground state properties, κ-(BEDT-TTF)$_2$X with X = Cu[N(CN)$_2$]Cl is a Mott insulator[10]. As t/U is increased, either by applying pressure or by changing the anion to X = Cu[N(CN)$_2$]Br (hereafter, κ-Br), there is an insulator-to-superconductor transition[11]. The X = Cu(NCS)$_2$ compound (hereafter κ-NCS), exhibiting a still larger t/U, remains a superconductor with a slightly lower transition temperature. The application of a hydrostatic pressure to κ-NCS increases t/U further and depresses T_c, suppressing superconductivity completely at ~4.5 kbar (ref. 9).

If pseudogap phenomena and superconducting fluctuations at temperatures well above T_c are in fact intrinsic properties of strongly correlated superconductors, rather than simply being artefacts of the complicated electronic environment found in the copper oxides, then the κ-(BEDT-TTF)$_2$X series should exhibit them too. Indeed, there are two key differences that favour the κ-(BEDT-TTF)$_2$X series as an experimental environment for studying superconductivity in the vicinity of a Mott transition. First, the parameter that determines the properties, t/U, is more amenable to theoretical analysis than doping and, furthermore, is not associated with the introduction of disorder into the system. Second, the upper critical magnetic fields of below 10 T and T_c of the order of 10 K make the entire phase diagram easily accessible under typical laboratory conditions, and mean that the normal state can be studied in a regime where thermal and magnetic energy scales are much smaller than the intrinsic energy scales, in particular the bandwidth.

Among the unusual phenomena exhibited by the copper oxides in the pseudogap phase, one of the most dramatic is the observation of a strongly enhanced Nernst effect above T_c (refs 12–15). The Nernst effect is the appearance of a transverse electric field E_y under the application of a temperature gradient $\nabla_x T$ in the presence of a perpendicular magnetic field B_z:

$$E_y = -e_N \nabla_x T = -N B_z \nabla_x T \quad (1)$$

where e_N is the Nernst signal, and N is the Nernst coefficient. In simple metals, the Nernst effect is generally small, linear in magnetic field (that is, N is constant), and has a sign that depends on the details of the bandstructure[15,16]. In contrast, a large Nernst effect is expected in the vortex liquid state: vortices can carry heat more effectively than the surrounding superconducting condensate, so they flow down an applied temperature gradient, causing an accumulation of phase in the condensate perpendicular to both the magnetic field and the temperature gradient, and hence a Nernst voltage. The Nernst effect has thus been identified as a sensitive indicator of the presence of superconducting fluctuations[12,17], and the onset of a substantial Nernst signal significantly above T_c in the underdoped copper oxides

[1]Clarendon Laboratory, Department of Physics, University of Oxford, OX1 3PU, UK. [2]Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA.

©2007 Nature Publishing Group

has been interpreted as strong evidence for the existence of superconducting fluctuations and vortex-like excitations in the pseudogap phase[12–15]. Nuclear magnetic resonance studies[7] indicate a reduction in the density of states above T_c in κ-(BEDT-TTF)$_2$X, particularly in the proximity of the Mott-insulating state, reminiscent of pseudogap-like behaviour. In this Letter we report a study of the superconducting fluctuations in the κ-(BEDT-TTF)$_2$X series via the Nernst effect[8,18].

Figure 2a shows the Nernst coefficient N for the large-t/U κ-NCS as a function of temperature for a range of magnetic fields applied perpendicular to the BEDT-TTF planes. Also shown (as a dashed line) is the resistance as a function of temperature in zero magnetic field, demonstrating a Fermi-liquid-like T^2 temperature dependence above the resistive transition at 9.8 K (ref. 19). In the normal state, the Nernst coefficient is small, negative and independent of magnetic field, as expected for a signal originating from quasiparticles. As the temperature drops through the superconducting transition and a vortex liquid state is established[20], the Nernst signal suddenly changes sign and increases dramatically in magnitude, peaking at about 7.5 K and then falling sharply as the irreversibility line is approached. Below 6 K, the mobility of the vortices decreases sharply as they freeze into a two-dimensional vortex lattice, and the Nernst signal drops. The resistive transition in a magnetic field (not shown here) occurs at the same temperature as the decline of the Nernst signal[21], because vortex flow contributes to a finite resistivity in the vortex liquid state.

In contrast to the linear normal state behaviour, N demonstrates a dramatic magnetic field dependence in the superconducting state, decreasing with increasing magnetic field. In this material[22] the penetration depth is $\lambda_{||} \approx 1\,\mu$m, while an applied magnetic field of 0.1 T corresponds to a vortex density of about 50 μm^{-2}. This implies that even at the lowest magnetic fields studied here the vortex dimensions are larger than the vortex separation. The repulsive interaction between vortices reduces the vortex mobility and this is manifested as a reduction of N as the vortex density increases; the vortex Nernst signal is almost completely suppressed by a magnetic field of 1.5 T. Thus the magnetic-field-dependent N is characteristic of a Nernst effect arising from vortices. Figure 2b shows a colour density plot of the Nernst signal e_N as a function of magnetic field and temperature; the vertical dashed line shows the zero-magnetic-field transition temperature. The contours are predominantly horizontal at temperatures above T_c, indicating that e_N is strongly dependent on magnetic field, but weakly dependent on temperature. The large signal within the superconducting state gives rise to predominantly vertical contours below T_c.

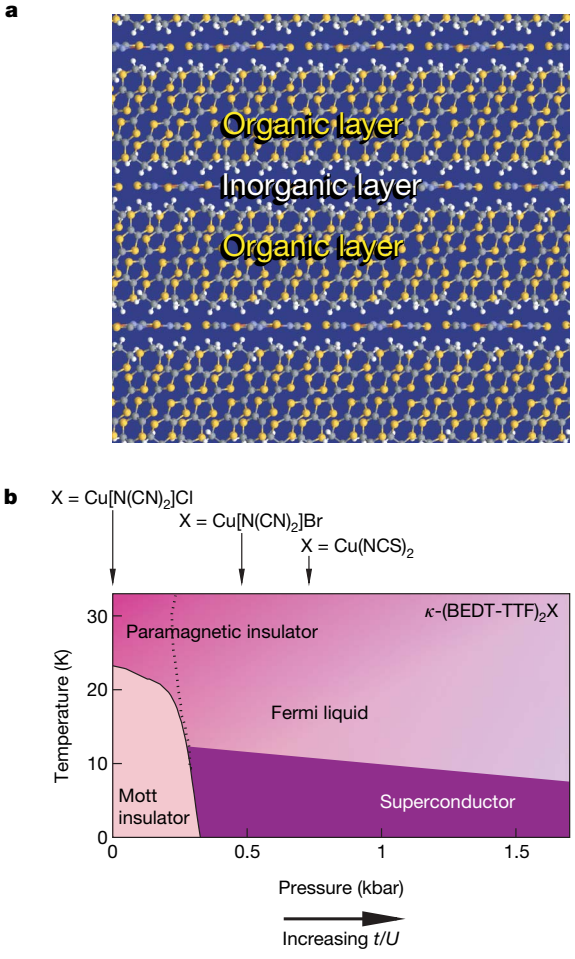


Figure 1 | Structure and phase diagram of κ-(BEDT-TTF)$_2$X. a, The crystal structure of the κ-(BEDT-TTF)$_2$X series of organic molecular conductors. The organic BEDT-TTF molecules stack in conducting planes separated by inorganic insulating layers, giving rise to a highly anisotropic quasi-two-dimensional band structure. **b**, The phase diagram as a function of temperature and effective pressure (adapted from figure 2 of ref. 8). Compounds with different X reside at different points on the horizontal axis at ambient pressure. Increasing pressure corresponds to increasing the ratio t/U (see text). At low t/U the ground state is a Mott insulator, becoming a paramagnetic insulator above about 30 K. For larger t/U, the ground state is superconducting, becoming a Fermi liquid above about 10 K.

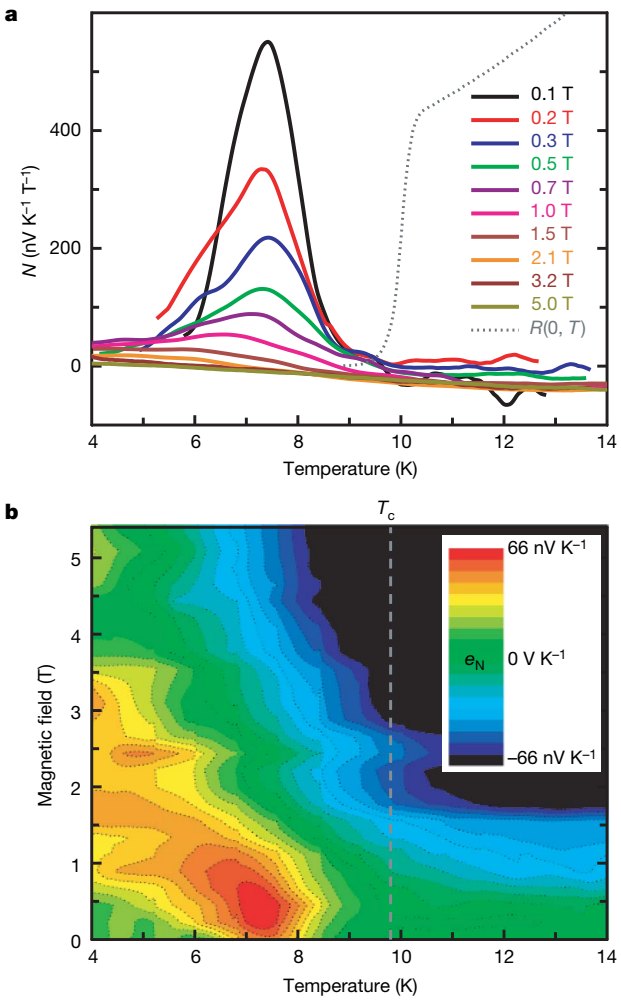


Figure 2 | The Nernst effect in κ-NCS. a, Solid lines: Nernst coefficient N as a function of temperature for a range of magnetic fields. The dashed line shows resistivity R as a function of temperature T in zero magnetic field. **b**, Colour density plot of Nernst signal e_N as a function of temperature and magnetic field. Above T_c, the Nernst signal is small, negative and linear in magnetic field consistent with a signal originating from quasiparticles. Below T_c, the Nernst signal is large, positive and nonlinear in magnetic field, consistent with a signal originating from flowing vortices.

©2007 Nature Publishing Group

There is nothing unexpected in these observations. The enhanced Nernst effect is observed only below the temperature at which a phase-coherent condensate is established and a superconducting transition is observed in the resistance. The phenomenology of the Nernst signal in both the normal and superconducting states agrees well with what is expected for a quasi-two-dimensional superconductor with a large t/U.

However, the Nernst effect in κ-Br, which is characterized by a smaller t/U (see Fig. 1), exhibits very unusual behaviour reminiscent of what has been observed in the underdoped copper oxides. Figure 3a shows N in this compound as a function of temperature for a range of magnetic fields, and (as a dashed line) the resistance in zero magnetic field indicating the superconducting transition at $T_c = 11.8$ K. Again, at high temperatures the Nernst coefficient is small, negative and independent of magnetic field, and below T_c there is a positive peak in N whose magnitude depends strongly on magnetic field and which appears at the temperature at which the vortices are most mobile. However, as the temperature falls through 18 K (50% above T_c), a positive and magnetic-field-dependent contribution to N begins to develop and the traces for different magnetic fields begin to diverge. This positive contribution develops smoothly into the peak within the superconducting state that is naturally attributed to vortices; there is no sharp change in behaviour at T_c. Figure 3b shows a colour density plot of the Nernst signal e_N as a function of magnetic field and temperature, with the zero-magnetic-field transition temperature marked by a vertical dashed line. This plot highlights the smooth onset of the positive Nernst signal above T_c, and a comparison with Fig. 2b

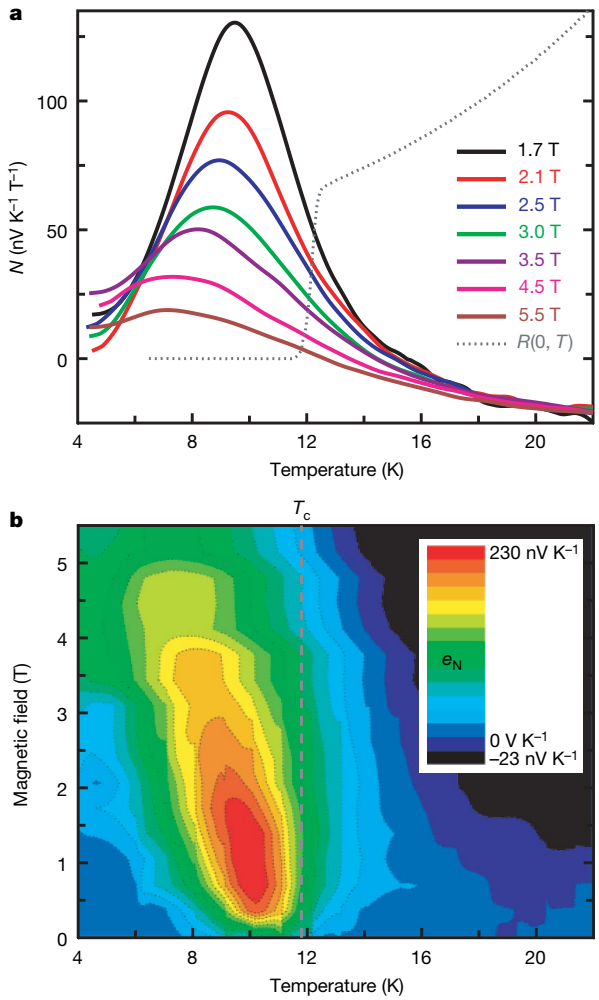


Figure 3 | The Nernst effect in κ-Br. **a**, As in Fig. 2a. **b**, As in Fig. 2b. The vortex-like signal emerges from the quasiparticle signal at temperatures exceeding T_c by 50%.

emphasizes the difference in the evolution of the Nernst signal between κ-NCS and κ-Br.

The behaviour of N between $T_c = 12$ K and 18 K in κ-Br cannot be explained as originating from normal-state quasiparticles, which generate a small, negative and magnetic-field-independent contribution. The smooth evolution of N into the vortex signal is a very strong indication that it is associated with the onset of superconducting fluctuations well above the temperature at which the phase-coherent condensate is established. This implies that we must consider two temperature scales separately: the temperature scale for the onset of superconducting fluctuations that are sufficiently robust to support a significant Nernst signal; and the (lower) temperature scale for the formation of a phase-coherent condensate supporting dissipationless electrical transport.

There are two main conclusions to be drawn. First, the phenomenology of our observations of superconducting fluctuations in κ-(BEDT-TTF)$_2$X is strikingly reminiscent of what has been reported in the copper oxide superconductors. Given that the two families of superconductors are so different structurally and electronically[4,8,23], we conclude that fluctuations of this kind are probably intrinsic to strongly correlated superconductors. Second, fluctuations are observed for small t/U but not for large t/U and we conclude that the proximity of the Mott transition is a key ingredient.

Kivelson and Emery[24] offered a very important insight in pointing out that superconductors with small superfluid density, and therefore low phase stiffness, should be susceptible to classical thermal fluctuations of the condensate phase. This mechanism even predicts the correct dependence for the importance of classical phase fluctuations in the copper oxides: as the doping goes down, the superfluid density decreases and classical phase fluctuations become more important. However, in traversing the κ-(BEDT-TTF)$_2$X series of superconductors via the parameter t/U instead of doping, the carrier density and Fermi surface parameters do not change dramatically[25], and in further contrast to the copper oxides, the superfluid density and critical temperature increase as the Mott state is approached[26–28]. Our observation that fluctuations become more important as t/U decreases is a further indication that we should be seeking a fluctuation mechanism that is associated with the proximity to the Mott insulating state.

In the Mott insulating state, the Coulomb energy is minimized by the occupation of each site by exactly one carrier (for a half-filled band). In contrast, the superconducting state is an example of a coherent state; the phase-coherence of the condensate is connected with an uncertainty in particle number[1]. When, within the superconducting state, the importance of Coulomb interactions is increased by reducing the ratio t/U, the number-uncertainty of the superconducting state incurs a Coulomb energy penalty[29]. This should increase the susceptibility of the condensate to phase fluctuations, and manifest itself macroscopically as a softening of the phase stiffness as the Mott-insulator transition is approached. (In the small t/U limit, the situation is treated by applying a Gutzwiller projection to remove any double-occupancy of sites from the coherent state[30]. Intermediate values of t/U may be handled by a partial projection[25,31,32].)

An enhancement of the Nernst signal may also be expected[17,33] to arise from gaussian fluctuations of the order parameter amplitude above T_c. This model[33] predicts comparable Nernst signals above T_c for the two compounds studied here, because their coherence lengths and conductivities are comparable. The dramatic difference that we actually observe therefore favours the phase fluctuation scenario described above, in which an enhancement of the Nernst signal as t/U decreases emerges naturally.

METHODS SUMMARY

Single crystals of κ-NCS and κ-Br, of typical dimensions $1 \times 1 \times 0.3$ mm^3 were synthesized electrochemically by the standard techniques[3]. Nine samples of κ-NCS and three samples of κ-Br were studied; there was no appreciable sample

©2007 Nature Publishing Group

dependence. The terminal ethylene groups of the BEDT-TTF molecule in κ-Br undergo an ordering transition at about 80 K. To ensure a minimum of ethylene-group disorder, the κ-Br samples were cooled at a maximum rate of 50 mK min^{-1} below 140 K. κ-NCS exhibits no such disorder and these samples were cooled at a typical rate of 200 mK min^{-1}.

A temperature gradient was applied by thermally coupling each end of the sample to a single-crystal quartz block using graphite paint and 50 μm gold wire. The quartz blocks were weakly thermally coupled to a copper plate; their temperature was controlled using resistive heaters and measured using low-magnetoresistance ruthenium oxide thermometers and thermocouples. The temperature difference applied was typically 250 mK. The experiment was maintained under a vacuum of below 10^{-5} mbar to ensure thermal isolation. The magnetic field was applied perpendicular to the BEDT-TTF conducting planes using a superconducting magnet in persistent mode to minimize electrical noise. Transverse thermal voltage contacts were made using resistive phosphor bronze wire of diameter 70 μm to reduce heat loss through these contacts. The signal was measured via oxygen-free copper wire with mechanical joints only from the low-temperature stage all the way to the Keithley voltmeter preamplifiers to minimize thermal voltage noise. The data were recorded over the range ~4 to ~35 K in temperature sweeps at a rate of 25 mK min^{-1} allowing sufficient averaging to limit the noise in the measurement to a few nanovolts.

Each temperature sweep was repeated for positive and negative magnetic fields. Longitudinal thermoelectric contributions to the voltage measured were removed by subtracting the positive and negative field sweeps.

Resistivity sweeps were made on the same samples using standard four-probe alternating-current (a.c.) lock-in techniques with a maximum current of 10 μA at a frequency of 87 Hz. The contacts used were the same as those used to perform the thermoelectric voltage measurements, which has the advantage that the sample need not be thermally cycled between thermoelectric and resistive measurements.

Received 4 June; accepted 15 August 2007.

1. Tinkham, M. *Introduction to Superconductivity* 2nd edn (Dover Publications, Mineola, New York, 2004).
2. Timusk, T. & Statt, B. The pseudogap in high-temperature superconductors: an experimental survey. *Rep. Prog. Phys.* **62,** 61–122 (1999).
3. Ishiguro, T., Yamaji, K. & Saito, G. *Organic Superconductors* 2nd edn (Springer, Berlin, 2006).
4. McKenzie, R. H. Similarities between organic and cuprate superconductors. *Science* **278,** 820–821 (1997).
5. Lefebvre, S. *et al.* Mott transition, antiferromagnetism, and unconventional superconductivity in layered organic superconductors. *Phys. Rev. Lett.* **85,** 5420–5423 (2000).
6. Kanoda, K. Metal–insulator transition in κ-$(ET)_2X$ and (DCNQI)2M: Two contrasting manifestation of electron correlation. *J. Phys. Soc. Jpn* **75,** 051007 (2006).
7. Mayaffre, H., Wzietek, P., Lenoir, C., Jerome, D. & Batail, P. C-13 study of a quasi-2-dimensional organic superconductor κ-$(ET)_2Cu[N(CN)_2]Br$. *Europhys. Lett.* **28,** 205–210 (1994).
8. Powell, B. J. & McKenzie, R. H. Strong electronic correlations in superconducting organic charge transfer salts. *J. Phys. Condens. Matter* **18,** R827–R866 (2006).
9. Caulfield, J. *et al.* Magnetotransport studies of the organic superconductor κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$ under pressure: the relationship between carrier effective mass and critical temperature. *J. Phys. Condens. Matter* **6,** 2911–2924 (1994).
10. Kagawa, F., Miyagawa, K. & Kanoda, K. Unconventional critical behaviour in a quasi-two-dimensional organic conductor. *Nature* **436,** 534–537 (2005).
11. Williams, J. M. *et al.* From semiconductor-semiconductor transition (42 K) to the highest T_c organic superconductor, κ-$(ET)_2Cu[N(CN)_2]Cl$ (T_c = 12.5 K)]. *Inorg. Chem.* **29,** 3272–3274 (1990).
12. Xu, Z. A., Ong, N. P., Wang, Y., Kakeshita, T. & Uchida, S. Vortex-like excitations and the onset of superconducting phase fluctuation in underdoped $La_{2-x}Sr_xCuO_4$. *Nature* **406,** 486–488 (2000).
13. Wang, Y. *et al.* High field phase diagram of cuprates derived from the Nernst effect. *Phys. Rev. Lett.* **88,** 257003 (2002).
14. Wang, Y. *et al.* Dependence of upper critical field and pairing strength on doping in cuprates. *Science* **299,** 86–89 (2003).
15. Wang, Y., Li, L. & Ong, N. P. Nernst effect in high-T_c superconductors. *Phys. Rev. B* **73,** 024510 (2006).
16. Sondheimer, E. H. The theory of the galvanomagnetic and thermomagnetic effects in metals. *Proc. R. Soc. Lond. Ser. A* **193,** 484–512 (1948).
17. Pourret, A. *et al.* Observation of the Nernst signal generated by fluctuating cooper pairs. *Nature Phys.* **2,** 683–686 (2006).
18. Logvenov, G. Y. *et al.* Anomalous Nernst effect in the mixed state of the two-band organic superconductors κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Br$ and κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$. *Physica C* **264,** 261–267 (1996).
19. Analytis, J. G. *et al.* Effect of irradiation-induced disorder on the conductivity and critical temperature of the organic superconductor κ-$(BEDT\text{-}TTF)_2Cu(SCN)_2$. *Phys. Rev. Lett.* **96,** 177002 (2006).
20. Sasaki, T. *et al.* Low-temperature vortex liquid states induced by quantum fluctuations in the quasi-two-dimensional organic superconductor κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$. *Phys. Rev. B* **66,** 224513 (2002).
21. Nam, M.-S. *et al.* Angle dependence of the upper critical field in the layered organic superconductor κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$ (BEDT-TTF ≡ bis(ethylene-dithio)tetrathiafulvalene). *J. Phys. Condens. Matter* **11,** L477–L484 (1999).
22. Carrington, A. *et al.* Low-temperature penetration depth of κ-$(ET)_2Cu[N(CN)_2]Br$ and κ-$(ET)_2Cu(NCS)_2$. *Phys. Rev. Lett.* **83,** 4172–4175 (1999).
23. Lee, P. A., Nagosa, N. & Wen, X.-G. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78,** 17–85 (2006).
24. Emery, V. J. & Kivelson, S. A. Importance of phase fluctuations in superconductors with small superfluid density. *Nature* **374,** 434–437 (1995).
25. Powell, B. J. & McKenzie, R. H. On the relationship between the critical temperature and the London penetration depth in layered organic superconductors. *J. Phys. Condens. Matter* **16,** L367–L373 (2004).
26. Lang, M., Toyota, N., Sasaki, T. & Sato, H. Magnetic penetration depth of κ-$(BEDT\text{-}TTF)_2Cu[N(CN)_2]Br$, determined from the reversible magnetization. *Phys. Rev. B* **46,** 5822–5825 (1992).
27. Larkin, M. I., Kinkhabwala, A., Uemura, Y. J., Sushko, Y. & Saito, G. Pressure dependence of the magnetic penetration depth in κ-$(BEDT\text{-}TTF)_2Cu(NCS)_2$. *Phys. Rev. B* **64,** 144514 (2001).
28. Pratt, F. L. & Blundell, S. J. Universal scaling relations in molecular superconductors. *Phys. Rev. Lett.* **94,** 097006 (2005).
29. Doniach, S. Quantum fluctuations in two-dimensional superconductors. *Phys. Rev. B* **24,** 5063–5070 (1981).
30. Anderson, P. W. The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235,** 1196–1198 (1987).
31. Powell, B. J. & McKenzie, R. H. Half-filled layered organic superconductors and the resonating-valence-bond theory of the Hubbard-Heisenberg model. *Phys. Rev. Lett.* **94,** 047004 (2005).
32. Gan, J. Y., Zhang, F. C. & Su, Z. B. Theory of gossamer and resonating valence bond superconductivity. *Phys. Rev. B* **71,** 014508 (2005).
33. Ussishkin, I., Sondhi, S. L. & Huse, D. A. Gaussian superconducting fluctuations, thermal transport, and the Nernst effect. *Phys. Rev. Lett.* **89,** 287001 (2002).

Acknowledgements Work at Oxford is funded by the EPSRC. Work at Argonne National Laboratory is supported by the Office of Basic Energy Sciences, Division of Materials Science, US Department of Energy. A.A. is supported by the Royal Society. We thank J. M. Bhaseen, K. Burnett, P. M. Chaikin, J. T. Chalker, L. Forro, D. Jaksch, N. P. Ong and I. A. Walmsley for discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.A. (arzhang.ardavan@physics.ox.ac.uk).

©2007 Nature Publishing Group

LETTERS

Single artificial-atom lasing

O. Astafiev[1,2], K. Inomata[2], A. O. Niskanen[3,4], T. Yamamoto[1,2,3], Yu. A. Pashkin[1,2], Y. Nakamura[1,2,3] & J. S. Tsai[1,2,3]

Solid-state superconducting circuits[1–3] are versatile systems in which quantum states can be engineered and controlled. Recent progress in this area has opened up exciting possibilities for exploring fundamental physics as well as applications in quantum information technology; in a series of experiments[4–8] it was shown that such circuits can be exploited to generate quantum optical phenomena, by designing superconducting elements as artificial atoms that are coupled coherently to the photon field of a resonator. Here we demonstrate a lasing effect with a single artificial atom—a Josephson-junction charge qubit[9]—embedded in a superconducting resonator. We make use of one of the properties of solid-state artificial atoms, namely that they are strongly and controllably coupled to the resonator modes. The device is essentially different from existing lasers and masers; one and the same artificial atom excited by current injection produces many photons.

Conventional lasers and masers consist of many atoms that are weakly coupled to a cavity owing to the tiny size of natural atoms. Nevertheless, by using a tightly confined cavity mode, coherent interaction of a single atom and the cavity can be achieved: the atom–cavity interaction time becomes shorter than the photon lifetime or the atom coherence time[10,11]. Such a strong coupling regime results in a qualitatively new feature: the vanishing of the pumping threshold that has been experimentally realized in single-atom masers and lasers[12,13]. On the other hand, quantum systems with artificial atoms allow one to easily make the interaction time much shorter than the coherence time, as has been demonstrated recently[4–6,14,15]. Furthermore, controllable interaction with a single cavity mode together with a fast mechanism of population inversion gives the possibility of realizing a lasing regime with many photons generated by one and the same atom[13,16,17].

In this work, we demonstrate lasing action in a maser based on a single Josephson-junction charge qubit with a population inversion mechanism provided by single-electron tunnelling events[18]. Alternative lasing schemes with superconducting qubits have been discussed elsewhere[19,20].

The artificial-atom maser consists of a resonator and a charge qubit coupled to it (Fig. 1). We fabricate a transmission-type half-wavelength coplanar-waveguide resonator using a 200-nm-thick Nb film (Fig. 1c). It has a bare resonance frequency $\omega_0/2\pi = 9.899$ GHz and a quality factor $Q = 7.6 \times 10^3$. The corresponding photon decay rate is $\kappa/2\pi = 1.3$ MHz. The qubit is fabricated by three-angle shadow evaporation of Al close to the end of the resonator, where the electric field is nearly maximal. The qubit[9] is well described by two charge states, $|0\rangle$ and $|2\rangle$, differing by one Cooper pair (consisting of two electrons) in the island, and is characterized by the Josephson energy E_J and the single-electron charging energy E_C. The electrostatic energy difference $\varepsilon = 4E_C(n_g - 1)$ between the two states is controlled by the normalized gate charge $n_g = C_g V_g/e$, where V_g is the gate voltage, C_g the gate capacitance and e the electron charge. The qubit eigenenergy follows $\Delta E = \sqrt{\varepsilon^2 + E_J^2}$ (top right panel of Fig. 1d). As the qubit is coupled to the resonator through the electric field ($\propto a^+ + a$, where a^+ and a are the photon creation and annihilation operators, respectively), the hamiltonian of the qubit-resonator system reads:

$$H = -\frac{1}{2}(\varepsilon\sigma_z + E_J\sigma_x) + \hbar\omega_0\left(a^+a + \frac{1}{2}\right) + \hbar g_0(a^+ + a)\sigma_z \quad (1)$$

The first term represents the qubit; σ_z and σ_x are the Pauli matrices. The second term describes the resonator. The interaction between the qubit and the resonator gives the third term, and is characterized by the coupling strength g_0. The value of $g_0/2\pi$ is found to be 80 MHz

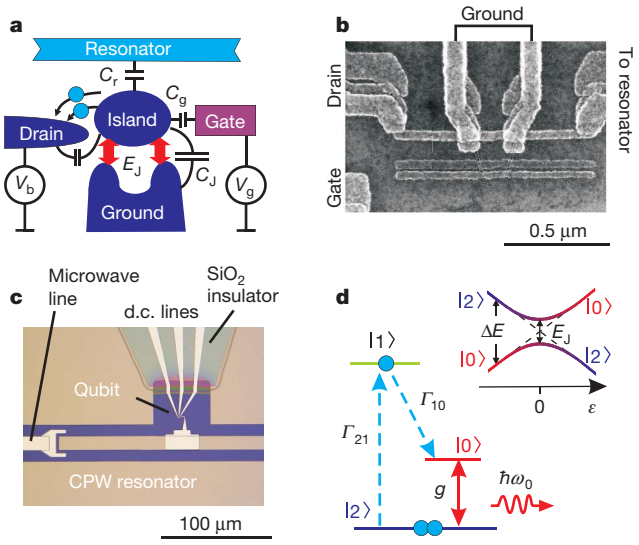


Figure 1 | Single artificial-atom maser and lasing mechanism. a, Schematic representation of the circuit. **b**, Scanning-electron micrograph of the qubit. The Josephson charge qubit consists of a superconducting Al island, with charging energy $E_C = e^2/2C_\Sigma = h$ 20 GHz (C_Σ is the total capacitance of the island), connected to ground through two Josephson junctions with a SQUID geometry so that the effective Josephson energy E_J is controlled by the magnetic flux Φ through the loop. A voltage-biased drain electrode is connected to the island via a 1.0-MΩ tunnel junction. A tail of an Al strip (see also **c**) forms another tunnel junction with estimated capacitance of $C_r \approx 200$ aF, defining the qubit–resonator coupling. The small junction conductance (~200 $k\Omega^{-1}$) is irrelevant for the unbiased junction. **c**, Micrograph of the left end of the Nb coplanar waveguide (CPW) resonator. At each end of the resonator, the central line is capacitively coupled to the external microwave line with a characteristic impedance of 50 Ω. The qubit is fabricated close to the end of the ~6.24-mm-long resonator. Bright stripes on top of the SiO_2 insulating layer are the qubit d.c. bias lines. An Al strip extends from the resonator towards the qubit for realizing strong capacitive coupling. **d**, Energy band diagram of the qubit (top right) and the lasing mechanism (bottom left). For $\varepsilon > 0$, population inversion is created by two sequential single-electron tunnelling events ($|2\rangle \rightarrow |1\rangle \rightarrow |0\rangle$) from the island to the drain.

[1]NEC Nano Electronics Research Laboratories, Tsukuba, Ibaraki 305-8501, Japan. [2]The Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198, Japan. [3]CREST-JST, Kawaguchi, Saitama 332-0012, Japan. [4]VTT Technical Research Center of Finland, Sensors, POB 1000, 02044 VTT, Espoo, Finland.

©2007 Nature Publishing Group

from fitting the dispersion curve observed in the transmission through the resonator when the qubit is biased at $\varepsilon = 0$ (Fig. 2a).

To create population inversion in the qubit, we introduce a drain electrode connected to the island via a tunnel junction with a resistance R_b of 1.0 MΩ (Fig. 1a, b). The drain electrode is voltage biased at a voltage V_b above $(2\Delta + E_C)/e$, which is required to extract two electrons from the island by breaking a Cooper pair (where Δ is the superconducting gap energy; $\Delta/h \approx 55$ GHz)[21,22]. As a result, $|2\rangle$ decays into $|0\rangle$ via two sequential single-electron tunnelling events in the incoherent process $|2\rangle \rightarrow |1\rangle \rightarrow |0\rangle$ with rates Γ_{21}, $\Gamma_{10} \approx (eV_b \pm E_C)/e^2R_b$ (positive sign for the former), respectively (bottom left panel of Fig. 1d). Therefore, the 'atom' is pumped into the $|0\rangle$ state regardless of the sign of ε. At $\varepsilon = 0$, a Cooper pair tunnels across the Josephson junction from the ground to the island ($|0\rangle \rightarrow |2\rangle$) without changing its energy. Thus, the so-called Josephson-quasiparticle (JQP) cycle involving the three charge states continues and results in a pronounced current peak[21,22]. For $\varepsilon \gg E_J$, the upper eigenstate of the qubit is nearly the $|0\rangle$ state, and the single-electron tunnelling process creates a population inversion with an effective rate $\gamma = \Gamma_{21}\Gamma_{10}/(\Gamma_{21} + \Gamma_{10})$. For $V_b = 0.65$ mV (used in the measurement below) $\gamma \approx 2.0 \times 10^9\ s^{-1}$ ($\gamma/2\pi \approx 320$ MHz), which is much larger than κ.

When ΔE is adjusted to $\hbar\omega_0$, the energy quantum of the qubit is transferred into the resonator as a photon, accompanied by a Cooper pair tunnelling across the Josephson junction ($|0, N\rangle \rightarrow |2, N+1\rangle$; here $|n, N\rangle$ represents a state with n excess electrons in the qubit island and N photons in the resonator). Completed by the pumping mechanism $|2, N+1\rangle \rightarrow |1, N+1\rangle \rightarrow |0, N+1\rangle$, the photon-assisted JQP

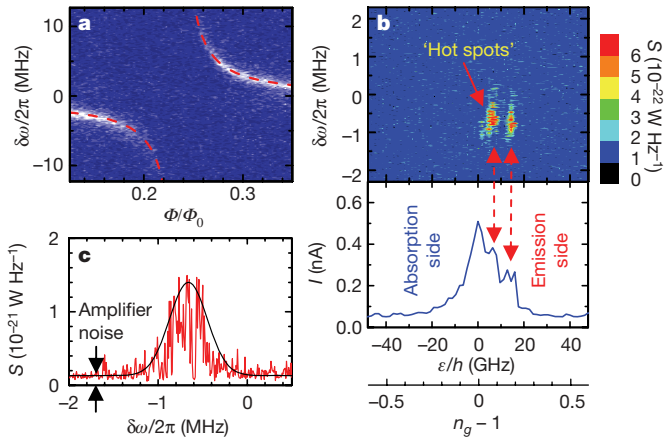


Figure 2 | Emission from the self-running maser. a, Transmission spectrum through the resonator measured with a weak microwave power ($P = -138$ dBm) as a function of the magnetic flux Φ in the SQUID loop. The average number of photons in the resonator is kept below 0.3. The detuning $\delta\omega/2\pi \equiv (\omega - \omega_0)/2\pi$ is the difference between the probe frequency $\omega/2\pi$ and the resonator frequency $\omega_0/2\pi$ (=9.899 GHz). The qubit is biased at $\varepsilon = 0$ and $V_b = 0$, so that the qubit energy ΔE equals E_J and there is no current injection. The observed dispersion curve is reproduced by $\delta\omega = \left((E_J - \hbar\omega_0) \pm \sqrt{(E_J - \hbar\omega_0)^2 + 4g_0^2}\right)/2\hbar$ with $E_J = E_{J0}\cos|\pi\Phi/\Phi_0|$ (where Φ_0 is the flux quantum), $E_{J0}/h = 13.7$ GHz and $g_0/h = 80$ MHz (red dashed lines). **b**, Emission power spectrum S from the resonator (upper panel) together with the current I through the qubit (lower panel) as a function of ε or n_g. The population inversion mechanism due to the JQP process is now activated with $V_b = 0.65$ mV. The Josephson energy of the qubit is reduced to $E_J/h = 5.4$ GHz by applying a magnetic flux $\Phi = 0.38\Phi_0$. The emission is seen as two 'hotspots', and the corresponding current peaks appear on the right slope of the JQP peak ($\varepsilon > 0$). This double hotspot feature is reproduced around every charge degeneracy point between $|n\rangle$ and $|n+2\rangle$, periodically in n_g. However, in another sample with lower γ, we observed a single hotspot with lower emission power. **c**, Emission power spectrum S at one of the hotspots taken at $\varepsilon/h = 7$ GHz (red curve). The black curve is an eye-guide envelope of the emission peak. The background level originates from the amplifier, which has a noise temperature of 10 K.

cycle proceeds repeatedly with increasing N, and N reaches a balance between photon generation and loss of the resonator. The coupling between $|0, N\rangle$ and $|2, N+1\rangle$ states is enhanced by a factor of $\sqrt{N+1}$; the photon field stimulates the photon generation process, which is analogous to stimulated emission in conventional lasers. However, in our case the photons are generated by one and the same atom. In conventional lasers, the ratio β (of spontaneous decay rate into the lasing mode to the total spontaneous decay rate) is very low. Therefore, a high pumping rate (above the lasing threshold) is required to achieve lasing. However in our system, with a single atom efficiently coupled to a single-mode cavity with β close to unity, the threshold no longer exists, and lasing takes place at any weak pumping rate[13,16].

In Fig. 2b, the emission power spectral density from the resonator (upper panel) is shown together with the current through the qubit (lower panel) as a function of ε. The observed current peak at $\varepsilon = 0$ is due to the JQP process. On the right-hand slope of the JQP peak ($\varepsilon > 0$; the emission side), two small current peaks ($I_p \approx 0.1$ nA above the JQP peak) appear. Correspondingly, we observe strong emission, shown as two 'hotspots' in the upper panel of Fig. 2b. The position of the first current peak and the hotspot corresponds to $\varepsilon/2\pi \approx 7 \pm 2$ GHz. Although the hotspot is rather broad, it is located consistently with the condition $\Delta E = \hbar\omega_0$ ($\varepsilon = 8.3$ GHz). (Although the emission takes place in a wide range of magnetic flux Φ, the data shown here are obtained at $\Phi = 0.38\Phi_0$ (Φ_0 is the flux quantum), where $E_J/h \approx 5.4$ GHz.) Because of finite ε, the effective coupling strength at the resonance is reduced to $g/2\pi = (g_0/2\pi)(E_J/\hbar\omega_0) \approx 44$ MHz.) One possible interpretation of the presence of the second hotspot is the two-photon resonance[18] expected at $\varepsilon = 19$ GHz. Note that the emission takes place only when the drain electrode is biased in the range $0.57\ \text{mV} \leq V_b \leq 0.71$ mV, where the JQP cycle is the dominant current carrying process. Note also that on the absorption side, $\varepsilon < 0$, microwave de-amplification is expected at $\Delta E = \hbar\omega_0$ and is indeed observed, though it is not shown here.

Figure 2c shows the emission spectrum at one of the hotspots. The frequency of the intense emission is shifted by about −0.7 MHz from the resonator frequency. The emission peak is unstable, showing low-frequency fluctuations, which can be attributed to the low-frequency charge noise. However, it is roughly confined within the envelope drawn by the black curve. The total emission power within the envelope is estimated to be $W = 7 \times 10^{-16}$ W, which corresponds to $N = 2(W/\hbar\omega_0)/\kappa \approx 30$ photons in the resonator. (The factor 2 comes from the equal probability that the photons escape from either end of the resonator: the number 30 may be underestimated, as the resonator internal loss is not accounted for.) The large number of photons accumulated in the single-mode resonator indicates a lasing effect, together with a linewidth narrower than both κ and Γ_{21}. However, the linewidth is still much wider than the quantum limit given by the Schawlow–Townes formula (see, for example, ref. 23) $\kappa/(2N)$ (of the order of $2\pi \times 10$ kHz), which means that it is broadened by some other mechanism, for example, charge fluctuations. The β-factor is estimated as a ratio of photon escape rate over the photon assisted Cooper-pair tunnelling rate $\beta > (N\kappa)/(I_p/2e) = 0.4$. It supports our picture of high lasing efficiency.

To additionally prove the lasing action of our device, we study the amplification of an external microwave signal. Figure 3a shows the normalized power and phase of coherent radiation output from the resonator. The blue curves show an ordinary transmission through the resonator when the qubit is biased away from the hotspots, and, as expected, the amplitude exhibits a lorentzian shape. The red curve demonstrates amplification of the drive microwave: at the hotspots, the transmission peak is enhanced on the low-frequency side of the bare resonant peak and slightly shifted towards lower frequencies with respect to the emission peak. At frequency $\delta\omega_{drive}/2\pi \approx -0.6$ MHz ($\delta\omega_{drive} \equiv \omega_{drive} - \omega_0$), the amplification switches to attenuation accompanied by the phase drop. One possible interpretation of the frequency shift and the phase

©2007 Nature Publishing Group

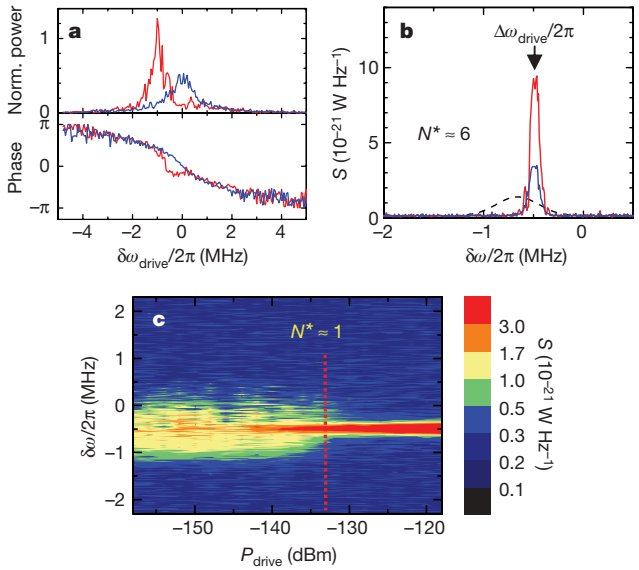


Figure 3 | Microwave amplification and injection locking. a, Normalized power (upper panel) and phase (lower panel) of the coherent radiation from the resonator as a function of detuning of the driving microwave $\delta\omega_{drive}/2\pi$, off the hot spots (blue, $\varepsilon/h = 40$ GHz) and at the hotspot (red, $\varepsilon/h \approx 7$ GHz) with $V_b = 0.65$ mV. The amplitude is normalized to the input driving power of $P_{drive} = -135$ dBm corresponding to $N^* = 0.6$ at $\delta\omega_{drive} = 0$. At $\delta\omega_{drive}/2\pi \approx -0.6$ MHz, the amplification regime switches to the attenuated transmission regime. The change of regime is also seen as a sudden phase drop. **b**, Output power spectrum S under the driving microwave field at a fixed detuning $\delta\omega_{drive}/2\pi = -0.5$ MHz and with a power $P_{drive} = -125$ dBm. The blue curve is measured outside hotspots, while the red curve is taken at the hotspot ($\varepsilon/h \approx 7$ GHz). The black dashed curve is the envelope of the emission spectrum in the absence of any microwave drive (see Fig. 2c). **c**, Output power spectrum S (colour map in log scale) as a function of the driving power for the detuning frequency $\delta\omega_{drive}/2\pi = -0.5$ MHz. The spectrum gets as narrow as the measurement bandwidth (100 kHz) when $N^* > \sim 1$.

drop in Figs 2c and 3a is that they are signatures of qubit–resonator coupling $g(N+1)^{1/2}$ that makes the system nonlinear in photon fields whose amplitude is proportional to $N^{1/2}$. The resonance frequency and consequently the amplification and emission peaks are expected to split by approximately $\pm g/2N^{1/2}$. The observed amplification peak is shifted by about -1 MHz, which is of the order of the expected value. However, the peaks shifted to positive frequencies are not observed. The phase shift accompanying an amplification peak of a narrow band amplifier should also be additionally affected by the nonlinear term $N^{-1/2}$ and therefore drops on the right-hand slope of the amplification peak, where N is suppressed.

Next we study the emission spectrum under the external driving microwave, expecting 'injection locking' effects[22]. The red curve in Fig. 3b exemplifies the emission power spectrum at the hotspot when the external drive power P_{drive} corresponding to six photons in the resonator ($N^* \equiv (P_{drive}/\hbar\omega_0)/\kappa \approx 6$) is applied. The driven emission (red curve) reproduces the shape of the drive signal (blue curve) at frequency $\omega_{drive}/2\pi$ ($\delta\omega_{drive}/2\pi = -0.5$ MHz), whereas the emission is suppressed at $\delta\omega \neq \delta\omega_{drive}$. This is consistent with the expected locking mechanism of the emission. The red peak is also much higher than the blue one, which is the transmitted spectrum measured with the same drive power and at $\varepsilon/h = 40$ GHz (outside the hotspots). We measured the injection locking in the range of $\delta\omega_{drive}/2\pi$ from -1.5 MHz to 0.5 MHz, and found that locking takes place at higher power for larger detuning from the emission peak maximum. The spectrum strongly depends on P_{drive} (Fig. 3c). When N^* exceeds 1, the emission line shrinks to the drive frequency with the width limited by the measurement bandwidth (100 kHz) and amplitude fluctuations in the locked signal are suppressed. The injection locking effect resulting in frequency stabilization and emission narrowing additionally proves the lasing action.

We have demonstrated a lasing effect in the simplest possible geometry—one 'atom' coupled to a resonator. The physical simplicity and controllability make it especially attractive for studying fundamental laser properties. We expect that the artificial-atom masers could be used as on-chip microwave sources and microwave amplifiers.

Received 15 March; accepted 31 July 2007.

1. Makhlin, Y., Schön, G. & Shnirman, A. Quantum-state engineering with Josephson-junction devices. *Rev. Mod. Phys.* **73,** 357–400 (2001).
2. Devoret, M. H., Wallraff, A. & Martinis, J. M. Superconducting qubits: A short review. Preprint at ⟨http://arxiv.org/abs/cond-mat/cond-mat/0411174⟩ (2004).
3. Wendin, G. & Shumeiko, V. S. in *Handbook of Theoretical and Computational Nanotechnology* Vol. 3 (eds Rieth, M. & Schommers, W.) 223–309 (American Scientific Publishers, Los Angeles, 2006).
4. Wallraff, A. *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431,** 162–167 (2004).
5. Chiorescu, I. *et al.* Coherent dynamics of a flux qubit coupled to a harmonic oscillator. *Nature* **431,** 159–161 (2004).
6. Johansson, J. *et al.* Vacuum Rabi oscillations in a macroscopic superconducting qubit LC oscillator system. *Phys. Rev. Lett.* **96,** 127006 (2006).
7. Schuster, D. I. *et al.* Resolving photon number states in a superconducting circuit. *Nature* **445,** 515–518 (2007).
8. Houck, A. A. *et al.* Generating single microwave photons in a circuit. Preprint at ⟨http://arxiv.org/abs/cond-mat/0702648⟩ (2007).
9. Nakamura, Y., Pashkin, YuA & Tsai, J. S. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398,** 786–788 (1999).
10. Raimond, J.-M., Brune, M. & Haroche, S. Manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.* **73,** 565–582 (2001).
11. Mabuchi, H. & Doherty, A. Cavity quantum electrodynamics: Coherence in context. *Science* **298,** 1372–1377 (2002).
12. Walther, H. *et al.* Cavity quantum electrodynamics. *Rep. Prog. Phys.* **69,** 1325–1382 (2006).
13. McKeever, J., Boca, A., Boozer, A. D., Buck, J. R. & Kimble, H. J. Experimental realization of a one-atom laser in the regime of strong coupling. *Nature* **425,** 268–271 (2003).
14. Reithmaier, J. P. *et al.* Strong coupling in a single quantum dot-semiconductor microcavity system. *Nature* **432,** 197–200 (2004).
15. Yoshie, T. *et al.* Vacuum Rabi splitting with a single quantum dot in a photonic crystal nanocavity. *Nature* **432,** 200–203 (2004).
16. Rice, P. R. & Carmichael, H. J. Photon statistics of a cavity-QED laser: A comment on the laser-phase transition analogy. *Phys. Rev. A* **50,** 4318–4329 (1994).
17. Mu, Y. & Savage, C. M. One-atom lasers. *Phys. Rev. A* **46,** 5944–5954 (1992).
18. Rodrigues, D. A., Imbers, J. & Armour, A. D. Quantum dynamics of a resonator driven by a superconducting single-electron transistor: A solid-state analogue of the micromaser. *Phys. Rev. Lett.* **98,** 067204 (2007).
19. You, J. Q., Liu, Y. X., Sun, C. P. & Nori, F. Persistent single-photon production by tunable on-chip micromaser with a superconducting qubit circuit. *Phys. Rev. B* **75,** 104516 (2007).
20. Hauss, J., Fedorov, A., Hutter, C., Shnirman, A. & Schön, G. Single-qubit lasing and cooling at the Rabi frequency. Preprint at ⟨http://arxiv.org/abs/cond-mat/0701041⟩ (2007).
21. Averin, D. V. & Aleshkin, V. Ya. Resonance tunneling of Cooper pairs in a system of small Josephson junctions. *JETP Lett.* **50,** 367–369 (1989).
22. Fulton, T. A., Gammel, P. L., Bishop, D. J., Dunkleberger, L. N. & Dolan, G. J. Observation of combined Josephson and charging effects in small tunnel junction circuits. *Phys. Rev. Lett.* **63,** 1307–1310 (1989).
23. Siegman, A. E. *Lasers* (University Science Books, Mill Valley, 1986).

Acknowledgements We are grateful to A. Zagoskin, A. Smirnov, L. Murokh, S. Kouno, A. Tomita and A. Clerk for discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to O.A. (astf@zb.jp.nec.com).

©2007 Nature Publishing Group

LETTERS

Persistence of full glacial conditions in the central Pacific until 15,000 years ago

P.-H. Blard[1,2]†, J. Lavé[3]†, R. Pik[2], P. Wagnon[4] & D. Bourlès[1]

The magnitude of atmospheric cooling during the Last Glacial Maximum and the timing of the transition into the current interglacial period remain poorly constrained in tropical regions, partly because of a lack of suitable climate records[1]. Glacial moraines provide a method of reconstructing past temperatures, but they are relatively rare in the tropics. Here we present a reconstruction of atmospheric temperatures in the central Pacific during the last deglaciation on the basis of cosmogenic ^{3}He ages of moraines and numerical modelling of the ice cap on Mauna Kea volcano, Hawaii—the only highland in the central Pacific on which moraines that formed during the last glacial period are preserved[2]. Our reconstruction indicates that the Last Glacial Maximum occurred between 19,000 and 16,000 years ago in this region and that temperatures at high elevations were about 7 °C lower than today during this interval. Glacial retreat began about 16,000 years ago, but temperatures were still about 6.5 °C lower than today until 15,000 years ago. When combined with estimates of sea surface temperatures in the central Pacific Ocean[3], our reconstruction indicates that the lapse rate during the Last Glacial Maximum was higher than at present, which is consistent with the proposal that the atmosphere was drier at that time[1,4]. Furthermore, the persistence of full glacial conditions until 15,000 years ago is consistent with the relatively late and abrupt transition to warmer temperatures in Greenland[5], indicating that there may have been an atmospheric teleconnection between the central Pacific and North Atlantic regions during the last deglaciation.

There is an important need for precise ice-retreat chronology in areas where past glaciers are the sole indicators of palaeoclimate. Mauna Kea volcano (elevation 4,206 m) is the only highland area of the central Pacific (Hawaii, 19° N) where moraines of the last glaciation have been preserved[2] (Fig. 1). Fortunately, many of the glacial deposits on Mauna Kea are not covered by volcanic material and have been exposed continuously since their formation. The emplacement of these moraines can therefore be dated by using cosmic-ray exposure methods. This late Pleistocene glacial footprint is represented by a ~10-km diameter cap of moraines draping the upper slopes of the volcano. These deposits extend downslope to a mean elevation of 3,400–3,500 m, with local outlet digitations reaching altitudes as low as 3,200 m in places such as the Pohakuloa gulch (Fig. 1c). The age of these glacial deposits has been

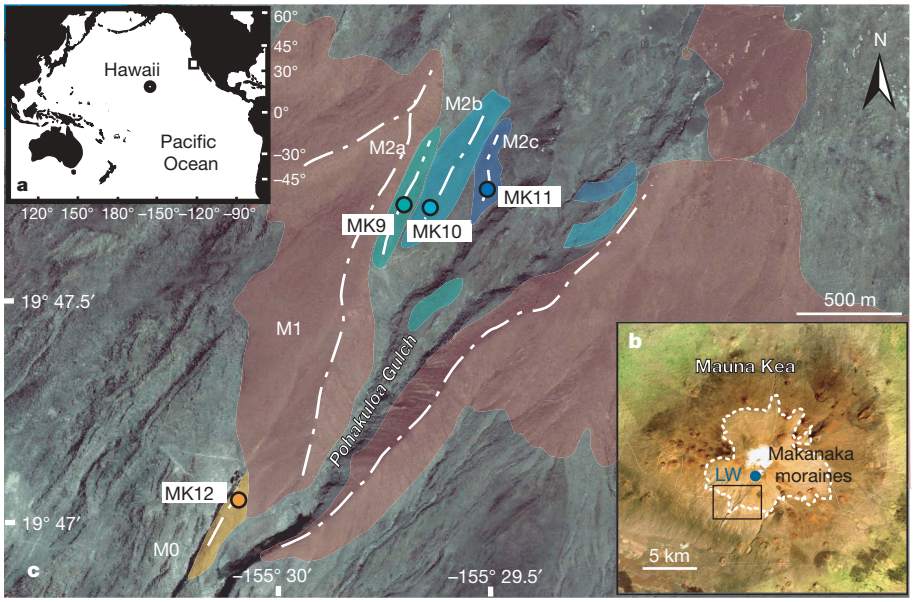


Figure 1 | Map of the sampled glacial deposits (Mauna Kea, Hawaii, central Pacific). a, Location of the Hawaiian Islands in the Pacific Ocean. The square shows the location of marine core MD02-2503, Santa Barbara Basin[27]. **b**, Location of the Pohakuloa gulch (SPOT image). LW, Lake Waiau. **c**, Map of moraines extents (M0 in orange, M1 in red, M2a, M2b and M2c in blue) and locations of the sampled boulders.

[1]Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement, CNRS – Aix Marseille Université, 13545 Aix en Provence, France. [2]Centre de Recherches Pétrographiques et Géochimiques, CNRS, 54501 Vandoeuvre-lès-Nancy, France. [3]Laboratoire de Géodynamique des Chaînes Alpines, CNRS – Université Joseph Fourier, 38400 Grenoble, France. [4]Institut de Recherche Pour le Développement, Great Ice, Laboratoire de Glaciologie et Géophysique de l'Environnement, 38402 Grenoble, France. †Present addresses: Geological and Planetary Science Division, California Institute of Technology, MC 100-23, 1200 E. California Boulevard, Pasadena, California 91125, USA (P.-H.B.); Centre de Recherches Pétrographiques et Géochimiques, CNRS, 54501 Vandoeuvre-lès-Nancy, France (J.L.).

©2007 Nature Publishing Group

roughly constrained between 31 ± 9 and 18 ± 10 kyr BP (errors are 1σ) by K–Ar dating of interbedded lavas[6]. Two cosmogenic ^{36}Cl ($^{36}Cl_c$) data on the main moraine M1 yielded ages of 18.9 ± 0.8 and 20.3 ± 2.3 kyr (ref. 7). The errors accompanying these ages, however, are analytical only and should include the ~30% uncertainties[8] in the production rate of $^{36}Cl_c$. More precise ages can be determined by using cosmogenic ^{3}He (^3He_c) because its production rate is known with better precision (1σ < 10%). However, measuring ^3He_c exposure ages requires the presence of the mineral olivine or pyroxene[9]. The only glacial deposits filling this petrological criterion were encountered within the Pohakuloa gulch (Fig. 1c). The shoulders dominating this talweg are the massive lateral moraine M1, whose frontal termination extends downslope to ~3,200 m, where it locally overlays frontal deposit M0 (Fig. 1c). M0 is interpreted as a subdued frontal moraine deposited in front of the Pohakuloa glacial tongue during its initial advance of the Last Glacial Maximum (LGM). M0 thus pre-dates the building of the main moraine M1. Three boulders ~1 m high (MK12A, MK12B and MK12C) have been sampled at ~3,200 m on M0 (Fig. 1c and Supplementary Table 1) to determine a maximum age for M1.

The retreat after the maximal ice extension is marked by three small recessional moraine deposits, M2a, M2b and M2c, extending from ~3,500 to 3,700 m within the inner part of the Pohakuloa gulch (Fig. 1c). To document the timing of glacial retreat, seven cobbles and boulders were sampled on M2a (MK9A, MK9B, MK9D and MK9F) and M2b (MK10A, MK10B and MK10C) at ~3,600 m, as well as the top of a large striated erratic block (MK11) in the M2c unit (Fig. 1c and Supplementary Table 1).

Measurement of cosmogenic ^{3}He concentrations is described in Methods. ^3He_c data were converted into exposure ages by using the ^3He_c production rate calibrated in Hawaii[10]. A sampling test performed on M2 led us to exclude boulders less than 80 cm in height, which improves the reliability of these cosmogenic ages (Supplementary Fig. 1).

The three boulders ~1 m high (MK12A, MK12B and MK12C) sampled in M0 have concordant ages of 17.2 ± 1.7, 19.1 ± 1.9 and 18.6 ± 1.9 kyr, respectively (Supplementary Table 1). The consistency of these ages suggests that post-depositional processes did not significantly affect the exposure history of the boulders. These data thus indicate a maximum age of ~19 kyr for the local LGM (LLGM) (Fig. 2a).

The largest boulders (more than 80 cm) from the M2 moraine yield ages ranging from 14.5 ± 1.5 to 15.7 ± 1.6 kyr. The clustering of these ages, as well as the small size of the M2 recessional moraines, suggest a short duration of the ~3,600-m stillstand at ~15.5 kyr BP. This interpretation is consistent with a $^{36}Cl_c$ exposure age of 14.7 ± 4.4 kyr from a boulder located above the M2 moraine[7]. The ^3He_c dates of M2 are also nearly synchronous with the calibrated[11] ^{14}C age of 14.7 ± 0.7 kyr from Lake Waiau[12] (3,962 m) that postdates the Mauna Kea ice cap.

We thus propose the following glacial chronology (Fig. 2a): first, initiation of the terminal moraine M0 deposition at 3,200 m by ~19 kyr BP, synchronous with or after the initial glacial advance; second, a major stillstand between 19 and 16 kyr BP, leading to the construction of the main moraine M1; third, a recessional phase at ~16 kyr BP with an episodic and short stillstand at ~15.5 kyr BP marked by the recessional moraines M2 at ~3,600 m; and fourth, a marked acceleration of retreat, leaving Mauna Kea almost completely free of ice after ~15 kyr BP. This chronology is consistent, within uncertainties, with the glacial records in the nearest glaciers of the Rocky Mountains (reviewed in ref. 13) and Mexico[14]. Dating of moraines in the tropical Andes[15] indicates an earlier local LGM there, suggesting different behaviour in the Southern Hemisphere.

To determine the high-altitude palaeoclimatic conditions in Hawaii since 19 kyr BP, the Mauna Kea ice cap has been reconstructed by using a numerical ice flux model[16] that combines both a mass-balance law and the influence of the local topography on glacier geometry (see legend to Fig. 3a). An input mass-balance curve (Fig. 3b) is calculated from a positive-degree-day (PDD) model[17] that includes temperature, precipitation, albedo and direct solar radiation (see Methods). Current temperatures are characterized by a seasonal amplitude of ~4 °C (Supplementary Fig. 3) and by a lapse rate of 5.4 ± 0.1 °C km^{-1} (Supplementary Fig. 4). The total annual precipitation decreases exponentially with elevation to a minimum of ~260 mm yr^{-1} above 4,000 m (Supplementary Fig. 5).

The presence of a glacier at 3,200 m in the Pohakuloa gulch requires lowering the equilibrium line altitude (ELA) to ~3,780 m during the LLGM (Fig. 3b). This shift is compatible with a former estimate based on the accumulation-area ratio method[18]. Assuming invariant precipitation in the PDD model, the depression in ELA requires a mean annual decrease in temperature of ~7 °C at 4,200 m. This estimate includes a correction of 1 °C, which takes into account the combined effects of subsidence and LLGM sea level lowering[19].

The sensitivity of the ELA to snowfall amount has been assessed by calculating all the (temperature, precipitation) solutions able to reproduce the decrease in ELA during the LLGM (Fig. 3c). This test shows that, in Hawaii, the ELA is more sensitive to air temperature

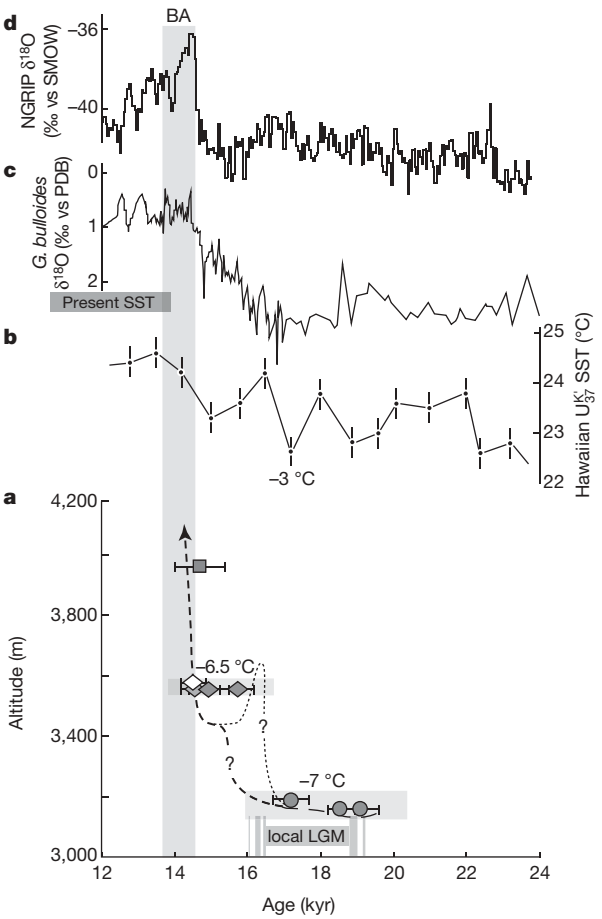


Figure 2 | Comparison of the Mauna Kea glacial chronology with other palaeoclimate proxies since 24 kyr BP. **a**, Glacial retreat chronology of the Pohakuloa gulch established by ^3He_c and ^{14}C dating[12] of moraines. Cosmogenic ages are plotted with their respective 1σ analytical uncertainties. For comparison with other absolute records, grey boxes include the 10% total uncertainty attached to production rates and scaling. Cooling values associated with each glacier position are computed for the corresponding ELA and assume invariant precipitations (Fig. 3). Grey square, ^{14}C age (ref. 12); diamonds, recessional moraines (light grey, M2a; dark grey, M2b; white, M2c); grey circles, terminal moraine M0; dashed lines, glacier front position. **b**, Hawaiian SST (PC17 core, 21° N, 158° W) from alkenone unsaturation index $U^{K'}_{37}$ (ref. 3). Error bars are shown as 1σ. **c**, δ^{18}O in *Globigerina bulloides* (SST proxy), Santa Barbara Basin (MD02-2503 core, 34° N, 120° W)[27]. **d**, $\delta^{18}O_{ice}$ (air temperature proxy) from the North Greenland Ice Core Project (NGRIP), Greenland (75° N, 42° W)[5]. BA = Bølling-Allerød.

©2007 Nature Publishing Group

than to precipitation variations. If, for example, precipitation is increased fivefold, a decrease in temperature of ~4.5 °C is still required to reproduce the LLGM snowline. Precipitation was probably lower during the LLGM because pollen records at 500 m elevation in Oahu Island (Hawaii) suggest that rainfall was 50–100% of present levels[20]. Thus, although some degree of decoupling may occur between low-altitude and high-altitude precipitation[21], our estimate of ~7 °C cooling at high altitudes could be considered to be a minimum value (Fig. 3c).

The accuracy of our estimate could be challenged on several counts. First, PDD models have been calibrated at high latitude and are not perfectly suited to describe outer-tropical glaciers in which daily cycles of melting are more sensitive to radiative budget than sensible heat flux[22]. However, melt data from glaciers in the tropical Andes[23] indicate that ablation is well correlated with annual variations in temperature. Moreover, the model used here was calibrated by using mass balance data from a well-documented tropical glacier (Zongo, 16° S) and includes the forcing of solar radiation and the influence of albedo (see Methods and Supplementary Information). Second, our model does not directly include effects of wind, cloudiness and air humidity, which are important in determining the mass balance of tropical glaciers and may therefore represent an important source of uncertainty. However, as reported[24], these effects can be integrated into the PDD model by varying the values of melting parameters. Thus, assessing the sensitivity of the model to extreme melting parameters (0.7-fold to 1.3-fold is the reported range for tropical glaciers[24]) provides an evaluation of the uncertainty arising from these atmospheric variables (±0.5 °C). Third, another source of uncertainty lies in the glacier dynamics model[16]. Sensitivity tests show that this imprecision is ±0.7 °C, which leads to a total uncertainty of ±0.9 °C for the model (see Supplementary Information). Last, if an extreme range in precipitation is considered (from half to double the present value[20,25]), these overall uncertainties imply that the LLGM cooling ranges between 8.3 and 5.4 °C (Fig. 3c).

This decrease in temperature is larger than the ~3.5 °C LGM cooling determined in ref. 21 at Mauna Kea. This 3.5 °C cooling was computed on the assumption that LGM precipitation (~1,600 mm yr^{-1}) were significantly above the present value of ~300 mm yr^{-1} at 3,800 m (ref. 21). With such a high snowfall, our model would yield a cooling of 4.4 ± 0.9 °C. The use of alternative precipitation may therefore partly account for the difference between both results. This previous estimate[21] also relies on a PDD model, and part of the discrepancy may also be due to the use of different melting parameters.

Coeval high-altitude ΔT of ~7 °C and local sea surface temperature (SST) cooling of ~3 °C during the LLGM[3] requires a steeper vertical gradient of surface temperature (~6.4 ± 0.2 °C km^{-1}, in contrast with the present value of 5.4 ± 0.1 °C km^{-1}). This result is qualitatively in agreement both with pollinic reconstructions[1,20] and with atmospheric general circulation models[4], indicating steeper lapse rates during the LLGM as a result of a drier tropical atmosphere. However, it should be noted that the glaciation of the Hawaiian summits possibly created a local feedback that might explain much of the high-atmosphere cooling of the area. This did not necessarily induce steeper free-air lapse rates in the whole Pacific Ocean.

If a constant lapse rate between 19 and 15 kyr BP is assumed, the glacial response shows variations in time and in temperature that are similar to those of the local Hawaiian SST[3,26] (Fig. 2a and b). Indeed, the ~0.7 °C sea surface warming between 17 and 15 kyr BP is similar to the ~0.5 °C increase necessary to trigger the increase in ELA from 3,780 to 3,880 m. Similarly, under invariant precipitation, the subsequent SST increase of 1.5 °C is sufficient to induce summit deglaciation after 15 kyr BP. Ice cap disappearance as a result of warming (rather than a decrease in precipitation) is also supported by pollinic reconstructions[20] showing that low-altitude precipitations

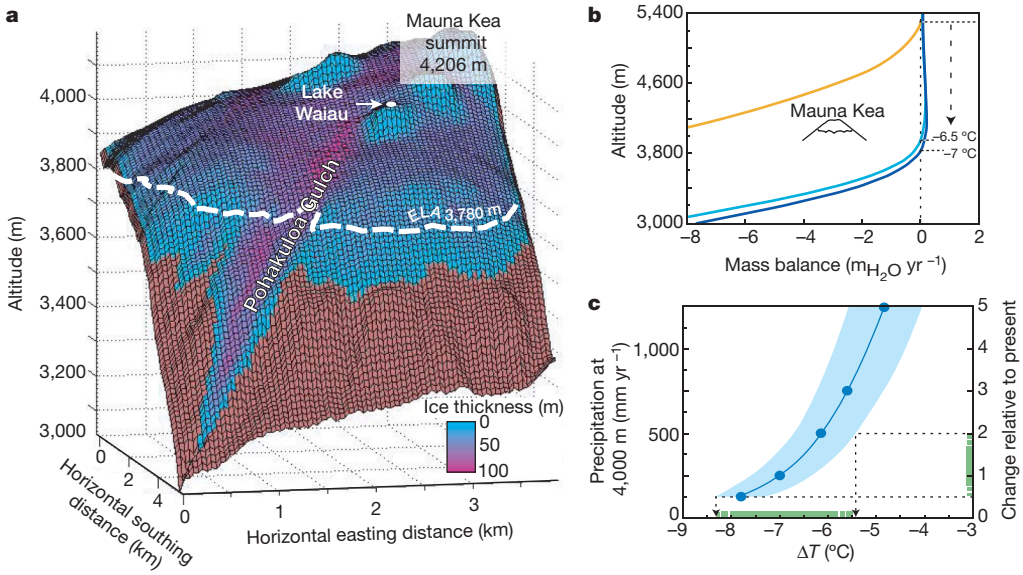


Figure 3 | Modelling of the Mauna Kea ice cap since the local Last Glacial Maximum, and palaeoclimatic reconstruction. a, Modelling of the maximal glacier extension in the Pohakuloa gulch (~3,200 m during the LLGM, 19–16 kyr BP) using the ice-flux model of ref. 16. Inputs are the 50-m-resolution digital elevation model of the Mauna Kea summit southwest area and the modelled elevation-dependent mass balance. Iterations on palaeoclimatic conditions were run until the modelled ice cap fitted the observed moraine extension. The best fit was obtained with an ELA 3,780 m high. **b**, Plot of present-day (orange), LLGM (16–19 kyr BP; dark blue) and 15 kyr BP (light blue) mass balance (MB) against elevation. MB calculations are described in Methods. The present-day Hawaiian MB curve is characterized by an ELA ~5,300 m high. The MB curve used to reproduce the LLGM extension (16–19 kyr BP) requires an ELA depression to ~3,780 m. The associated cooling is ~7 °C, under invariant precipitation. The MB corresponding to the 15 kyr BP extension yields an ELA at 3,880 m and a respective cooling of ~6.5 °C. **c**, Palaeoclimatic conditions able to reproduce the glacier extent during the local LGM (between 19 and 16 kyr BP). This plot of the change in temperature (ΔT) above 3,000 m against the change in precipitation at 4,000 m provides a direct test of the sensitivity of cooling to variations in precipitation. These climatic reconstructions are corrected for subsidence and change in sea level[19]. The dark blue line is calculated by using the model parameters calibrated from the Zongo glacier data (see Methods and Supplementary Information), the blue shading shows the total uncertainty of the model (see Supplementary Information) and the green box shows the conservative temperature interval on the assumption of a precipitation range during the LLGM of between half and double the present value[20,25].

©2007 Nature Publishing Group

did not decline, but actually increased by ~50% after 16 kyr BP. Even though uncertainties in timing preclude a definitive demonstration, this temperature coupling could indicate a local SST forcing of the snowline[4]. The second nearest marine core, from Santa Barbara Basin[27], provides time variations of $\delta^{18}O$ in *Globigerina bulloides* (SST proxy) showing a similar pattern (Fig. 2c). This eastern Pacific area began to warm at ~16.5 kyr BP, concurrently with the end of the Mauna Kea LLGM. Final ocean warming, after a slight cooling episode at ~15 kyr BP, was also nearly synchronous, within uncertainties, with the Hawaiian glacial retreat.

Finally, the proposed climate modelling requires the persistence of a 6.5 ± 0.9 °C cooler atmosphere until 15 kyr BP (Fig. 3b). This suggests a delayed warming of the central Pacific, whereas the southern Pacific[28] and Antarctic[29] had warmed continuously since 18 kyr BP. This late warming is concomitant with the Bølling–Allerød event and is therefore consistent with the Greenland pattern[5] (Fig. 2d). Thus, even if partly controlled by changes in SST, Hawaiian ELA fluctuations could also have been driven by an atmospheric teleconnection between the central Pacific and the northern Atlantic. Such atmosphere forcing is also supported by modelling experiments testing the sensitivity of the response of the northern Pacific to changes in the North Atlantic[30].

METHODS SUMMARY

Cosmogenic ^{3}He measurements were performed at the Centre de Recherches Pétrographiques et Géochimiques (Nancy) noble gases laboratory in accordance with the analytical procedure described previously[10]. Three samples of pure 0.5–1-mm phenocrysts were first slightly crushed *in vacuo* for evaluation of the $^3He/^4He$ magmatic ratio. To prevent any underestimate of the matrix-sited cosmogenic ^{3}He (ref. 10), the finest fraction (less than 0.15 mm) was removed before the subsequent *in vacuo* melting step. Because the contribution of magmatic helium is very low for these phenocrysts, some samples (the 0.5–1-mm fraction) were directly melted without preliminary crushing.

^3He_c concentrations (Supplementary Table 1) were calculated by assuming, first, that all ^{4}He extracted by melting was magmatic and second, that the magmatic $^3He/^4He$ ratio was 8 ± 1 Ra (where $Ra = 1.384 \times 10^{-6}$ is the atmospheric $^3He/^4He$ ratio). Post-eruptive addition of radiogenic $^4He^*$ is negligible (less than 5×10^{-15} mol g^{-1}) as a result of the low U and Th concentrations (~0.3 and 0.8 p.p.m., respectively) of these ~100-kyr-old basalts[6]. Measured ^3He_c concentrations range from 8.0×10^6 to 15.5×10^6 atoms g^{-1} and have very low analytical uncertainties ($1\sigma \approx 3\%$). Such precision was possible because the analysed phenocrysts are particularly poor in magmatic helium ($^4He < 5 \times 10^{-14}$ mol g^{-1}), and thus the correction for primordial ^{3}He does not exceed 3%. ^3He_c concentrations were converted into exposure ages by using a sea-level high-latitude production rate of 128 ± 5 atoms g^{-1} yr^{-1} (ref. 10). Geometric and time corrections of ^3He_c production rates were performed in accordance with the procedure described in Methods.

The Hawaiian palaeoclimate was determined by reconstructing past glacial extents with the use of a numerical model (Fig. 3). For this purpose, an ice-flow cellular automaton[16] was implemented with an ice mass-balance model based on a PDD approach. This PDD model includes the past variations of solar radiations[17] and the influence of albedo. The snow-rain threshold is +1 °C.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 August 2006; accepted 1 August 2007.

1. Farrera, I. *et al.* Tropical climates at the Last Glacial Maximum: a new synthesis of terrestrial palaeoclimate data. I. Vegetation, lake levels and geochemistry. *Clim. Dyn.* **15,** 823–856 (1999).
2. Porter, S. C. Hawaiian glacial ages. *Quat. Res.* **12,** 161–187 (1979).
3. Lee, K. E., Slowey, N. C. & Herbert, T. D. Glacial sea surface temperatures in the subtropical North Pacific: A comparison of U-K_{37}', $\delta^{18}O$, and foraminiferal assemblage temperature estimates. *Paleoceanography* **16,** 268–279 (2001).
4. Kageyama, M., Harrison, S. P. & Abe-Ouchi, A. The depression of tropical snowlines at the last glacial maximum: What can we learn from climate model experiments? *Quaternary Int.* **138,** 202–219 (2005).
5. Andersen, K. K. *et al.* High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431,** 147–151 (2004).
6. Wolfe, E. W., Wise, W. S. & Dalrymple, G. B. The geology and petrology of Mauna Kea volcano, Hawaii—a study of postshield volcanism. *Prof. Pap. US Geol. Surv.* 1557 (1997).
7. Dorn, R. I. *et al.* Glacial chronology. *Res. Explor.* **7,** 456–471 (1991).
8. Gosse, J. C. & Phillips, F. M. Terrestrial *in situ* cosmogenic nuclides: theory and application. *Quat. Sci. Rev.* **20,** 1475–1560 (2001).
9. Trull, T. W. & Kurz, M. D. Experimental measurements of ^{3}He and ^{4}He mobility in olivine and clinopyroxene at magmatic temperatures. *Geochim. Cosmochim. Acta* **57,** 1313–1324 (1993).
10. Blard, P.-H. *et al.* Cosmogenic ^{3}He production rates revisited from evidences of grain size dependent release of matrix sited helium. *Earth Planet. Sci. Lett.* **247,** 222–234 (2006).
11. Reimer, P. J. *et al.* IntCal04 terrestrial radiocarbon age calibration, 0–26 cal kyr BP. *Radiocarbon* **46,** 1029–1058 (2004).
12. Peng, L. & King, J. W. A late quaternary geomagnetic secular variation record from Lake Waiau, Hawaii, and the question of the Pacific nondipole low. *J. Geophys. Res. Solid Earth* **97,** 4407–4424 (1992).
13. Schaefer, J. M. *et al.* Near-synchronous interhemispheric termination of the last glacial maximum in mid-latitudes. *Science* **312,** 1510–1513 (2006).
14. Vázquez-Selem, L. & Phillips, F. M. in *Program and Abstracts of the 15th Biennial Meeting, American Quaternary Association, AMQUA 1998, Northern Hemisphere-Southern Hemisphere Interconnections (5–7 September 1998; Puerto Vallarta, México)*, p. 174.
15. Smith, J. A., Seltzer, G. O., Farber, D. L., Rodbell, D. T. & Finkel, R. C. Early local last glacial maximum in the tropical Andes. *Science* **308,** 678–681 (2005).
16. Harper, J. T. & Humphrey, N. F. High altitude Himalayan climate inferred from glacial ice flux. *Geophys. Res. Lett.* **30,** 1764–1767 (2003).
17. Hock, R. A distributed temperature-index ice and snowmelt model including potential direct solar radiation. *J. Glaciol.* **45,** 101–111 (1999).
18. Porter, S. C. Pleistocene snowlines and glaciation of the Hawaiian Islands. *Quaternary Int.* **138,** 118–128 (2005).
19. Ludwig, K. R., Szabo, B. J., Moore, J. G. & Simmons, K. R. Crustal subsidence rate off Hawaii determined from $^{234}U/^{238}U$ ages of drowned coral reefs. *Geology* **19,** 171–174 (1991).
20. Hotchkiss, S. & Juvik, J. O. A. Late-Quaternary pollen record from Ka'au crater, O'ahu, Hawai'i. *Quat. Res.* **52,** 115–128 (1999).
21. Hostetler, S. W. & Clark, P. U. Tropical climate at the last glacial maximum inferred from glacier mass-balance modeling. *Science* **290,** 1747–1750 (2000).
22. Sicart, J.-E., Wagnon, P. & Ribstein, P. Atmospheric controls of the heat balance of Zongo Glacier (16°S, Bolivia). *J. Geophys. Res. Atmos.* **110,** doi:10.1029/2004JD005732 (2005).
23. Francou, B., Vuille, M., Wagnon, P., Mendoza, J. & Sicart, J. E. Tropical climate change recorded by a glacier in the central Andes during the last decades of the twentieth century: Chacaltaya, Bolivia, 16°S. *J. Geophys. Res. Atmos.* **108,** doi:10.1029/2002JD002959 (2003).
24. Hock, R. Temperature index melt modelling in mountain areas. *J. Hydrol.* **282,** 104–115 (2003).
25. Kitoh, A., Murakami, S. & Koide, H. A simulation of the last glacial maximum with a coupled atmosphere–ocean GCM. *Geophys. Res. Lett.* **28,** 2221–2224 (2001).
26. Lee, K. E. & Slowey, N. C. Cool surface waters of the subtropical North Pacific Ocean during the last glacial. *Nature* **397,** 512–514 (1999).
27. Hill, T. M. *et al.* Pre-Bølling warming in Santa Barbara Basin, California: surface and intermediate water records of early deglacial warmth. *Quat. Sci. Rev.* **25,** 2835–2845 (2006).
28. Kiefer, T. & Kienast, M. Patterns of deglacial worming in the Pacific Ocean: a review with emphasis on the time interval of Heinrich event 1. *Quat. Sci. Rev.* **24,** 1063–1081 (2005).
29. Johnsen, S. J., Dansgaard, W., Clausen, H. B. & Langway, C. C. Jr. Oxygen isotope profiles through the Antarctic and Greenland ice sheets. *Nature* **235,** 429–434 (1972).
30. Mikolajewicz, U., Crowley, T. J., Schiller, A. & Voss, R. Modelling teleconnections between the North Atlantic and North Pacific during the Younger Dryas. *Nature* **387,** 384–387 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Bard and S. Kidder for advice on the manuscript; S. Rowland and F. Trusdell for their assistance in the field; N. Humphrey for sharing his ice-flux Matlab code; G. Leduc, C. Vincent, R. Hock, D. Paillard, N. Thouveny, S. Sépulcre and G. Brocard for discussions that helped to improve glacial modelling and palaeoclimatic interpretations; and L. Zimmerman and B. Tibari for their analytical assistance in the Centre de Recherches Pétrographiques et Géochimiques (CRPG) noble gases laboratory. We thank the State of Hawaii for delivering sampling permits. Financial support was provided by the French INSU programme 'Relief de la Terre'.

Author Contributions P.-H.B. and J.L. conducted the field work in Hawaii, numerical modelling, data interpretation and paper writing. P.-H.B. and R.P. performed the cosmogenic ^{3}He analyses at CRPG (Nancy). P.W. provided ablation and climatic data from the Zongo glacier and helped in developing the glacier mass-balance model. D.B. participated in interpreting the cosmogenic data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.-H.B. (blard@gps.caltech.edu or phb26@yahoo.fr).

©2007 Nature Publishing Group

METHODS

Determination of cosmogenic ^{3}He ages. ^{3}He cosmogenic concentrations (Supplementary Table 1) were converted into exposure ages, using a sea-level high-latitude production rate of 128 ± 5 atoms g^{-1} yr^{-1}. This empirical value was obtained from olivines with similar compositions to those analysed here[10]. The use of the same reference rate for pyroxenes is justified because ^3He_c concentrations are undistinguishable, within uncertainties, for both phenocrysts species coexisting within the same objects (MK9D and MK11, Supplementary Table 1). Ages were corrected for mask topography (less than 3%)[31] and boulder geometry (correction between 5% and 7%)[32]. Spatial and air-pressure corrections were performed by using the scaling factors of Stone[33], assuming that the island has been affected by a constant subsidence rate of 2.6 ± 0.4 mm kyr^{-1} (ref. 19). Effects of geomagnetic fluctuations on production rates were taken into account by using the corrections of Dunai[34] and the magnetic database of Carcaillet *et al.*[35]. These magnetic corrections are 6% and 3% for M2 and M0, respectively.

Glacier modelling. The cellular automata developed in ref. 16 was used to model the stationary state of the past glacial extents in Mauna Kea (details are given in Supplementary Information). This code was implemented with the mass-balance model described below.

'PDD-solar radiation' mass-balance model. The glacier mass balance was determined using an improved PDD model. This model, which integrates not only the pure temperature forcing but also the effect of solar radiation, has demonstrated its ability in reproducing the observed mass balance at high spatial and time resolutions[17]. Because the purpose of this study is to interpret the palaeoglacier in terms of mean palaeoclimatic variables, the mass balance was calculated over a one-year cycle, as follows.

The annual snow accumulation S (mm yr^{-1}) is the sum of the monthly snowfalls:

$$S = \sum_{i=1}^{12} S_i \tag{1}$$

S_i (mm per month) is calculated for each month assuming that precipitation falls as snow when the temperature is below the rain–snow threshold T_S. The PDD model assumes that temperatures have a normal distribution around the monthly mean temperature T_{mi}, with a standard deviation σ_i:

$$S_i = P_i \frac{1}{\sigma_i \sqrt{2\pi}} \int_{-\infty}^{T_S} e^{\frac{-(T-T_{mi})^2}{2\sigma_i^2}} dT \tag{2}$$

where P_i (mm per month) is the monthly precipitation. We assumed that T_S is 1 °C, which is a value reported from the tropical Andes[36]. Data do indeed show that below 0 °C 100% of precipitation is snow, and that above 2 °C 100% is rainfall. Moreover, weather stations of the Mauna Kea summit indicate that annual snowfalls are 125 ± 25 mm water equivalent, which is consistent with the value of 120 mm yr^{-1} predicted by our model and thus justifies the use of 1 °C for T_S. However, for Hawaii this parameter has a very limited influence on the climatic conclusions because Mauna Kea precipitations are particularly low (~250 mm yr^{-1} at 4,200 m). Indeed, a variation of ± 2 °C in T_S would induce an ELA shift of less than 50 m.

To consider the influence of albedo, which implies that the melting rate of ice is higher than that of snow (ice being darker than snow), we used different ablation parameters for snow and ice. The monthly ablation A_i (mm yr^{-1}) was calculated by assuming that the total melting is the sum of snow ablation A_{snow} and ice ablation A_{ice} (when ablation is sufficient to start melting the ice):

$$A_i = A_{snow}\,(i) + A_{ice}\,(i) \tag{3}$$

Given that α is the fraction of time during which ablation affects only snow, $\alpha = \frac{A_{snow}(i)}{A_i}$, we can write

$$A_{snow}(i) = \alpha \frac{365}{12} (M_{snow} + a_{snow} I_i) T_{Mi} \tag{4}$$

$$A_{ice}(i) = (1-\alpha) \frac{365}{12} (M_{ice} + a_{ice} I_i) T_{Mi} \tag{5}$$

where M_{snow} and M_{ice} (mm d^{-1} °C^{-1}) are the melt factors of snow and ice, respectively, a_{snow} and a_{ice} (mm d^{-1} W^{-1} m^2 °C^{-1}) are the radiation melt factors of snow and ice, respectively, I_i (W m^{-2}) is the mean monthly solar radiation, and T_{Mi} (°C) is the integrated monthly positive temperature.

Because, for each month, the amount of total snow input is known (S_i), equation (4) can be used to calculate α:

$$\alpha = \frac{\frac{12}{365} S_i}{(M_{snow} + a_{snow} I_i) T_{Mi}} \tag{6}$$

If $\alpha < 1$, the monthly total ablation A_i is not only sufficient to remove all the snow cover but it also starts melting the glacier ice. Then, in this case, A_i is calculated as

$$A_i = S_i + (1-\alpha) \frac{365}{12} (M_{ice} + a_{ice} I_i) T_{Mi} \tag{7}$$

In contrast, if $\alpha > 1$, the snow accumulation is higher than the total ablation; in this case, only snow melting occurs and monthly ablation A_i is

$$A_{snow}(i) = \frac{365}{12} (M_{snow} + a_{snow} I_i) T_{Mi} \tag{8}$$

Then the annual ablation A (mm yr^{-1}) is obtained by summing the monthly ablations:

$$A = \sum_{i=1}^{12} A_i \tag{9}$$

Past values of the solar radiation I_i are integrated over the appropriate period (19–16 kyr BP and 15 kyr BP for Hawaii) by using the astronomical corrections of ref. 37 and the numerical solution of ref. 38. At each grid cell, I_i is also corrected for the landscape characteristics, with

$$I_i = I_0 \left(\frac{R_m}{R_i}\right)^2 \psi^{\left(\frac{P}{P_0 \cos Zi}\right)} \cos\theta \tag{10}$$

where I_0 is the solar constant (1,368 W m^{-2}), R_i is the mean monthly Sun–Earth distance, R_m the mean Sun–Earth distance, $\psi = 0.75$ is the atmospheric clear-sky transmissivity, P and P_0 are the local and sea-level atmospheric pressure, respectively, Z_i is the mean monthly zenith angle and θ is the angle of incidence between the normal to the grid slope and the solar beams. For Hawaii, P and P_0 are from the daily radiosonde data of the Hilo International Airport over the period 1998–2002 (ref. 39).

The integrated monthly positive temperature T_{Mi} is computed with the PDD approach[40], which assumes that temperatures have a normal distribution around the monthly mean temperature T_{mi}, σ_i being the monthly standard deviation:

$$T_{Mi} = \frac{1}{\sigma_i \sqrt{2\pi}} \int_0^{\infty} T e^{\frac{-(T-T_{mi})^2}{2\sigma_i^2}} dT \tag{11}$$

Annual accumulation S and ablation A are calculated at each elevation by using equations (1) to (11). Finally, the mass balance elevation law, B_m (mm yr^{-1}) is computed by subtracting ablation from accumulation:

$$B_m = S - A \tag{12}$$

Calibration of the model parameters. Several studies[17,41] have reported empirical values for the parameters M and a, but these studies were conducted on Scandinavian glaciers. Therefore, to avoid any bias due to the tropical and high mountain specificities of Hawaii, we used the monitored mass balance data from a well-studied tropical glacier (Zongo, Bolivia, 16° S) to tune the parameters M_{snow}, M_{ice}, a_{snow} and a_{ice} (Supplementary Fig. 2).

The best fit yielded the following values:

$$M_{snow} = 3.7 \text{ mm d}^{-1} \text{ °C}^{-1}$$

$$M_{ice} = 7.4 \text{ mm d}^{-1} \text{ °C}^{-1}$$

$$a_{snow} = 2.8 \times 10^{-3} \text{ mm d}^{-1} \text{ W}^{-1} \text{ m}^2 \text{ °C}^{-1}$$

$$a_{ice} = 5.3 \times 10^{-3} \text{ mm d}^{-1} \text{ W}^{-1} \text{ m}^2 \text{ °C}^{-1}$$

These parameters are not significantly different from the empirical values determined in ref. 41. Moreover, the snow parameters are lower than those of ice, which suggests that this empirical model properly integrates the effects of albedo. After this preliminary calibration, the model was applied to the Mauna Kea volcano (19° N) with the climatic data recorded by the Hawaiian weather stations (Supplementary Figs 3–5).

©2007 Nature Publishing Group

31. Dunne, J., Elmore, D. & Muzikar, P. Scaling factors for the rates of production of cosmogenic nuclides for geometric shielding and attenuation at depth on sloped surfaces. *Geomorphology* **27,** 3–11 (1999).
32. Masarik, J. & Wieler, R. Production rates of cosmogenic nuclides in boulders. *Earth Planet. Sci. Lett.* **216,** 201–208 (2003).
33. Stone, J. O. Air pressure and cosmogenic isotope production. *J. Geophys. Res. Solid Earth* **105,** 23753–23759 (2000).
34. Dunai, T. J. Influence of secular variation of the geomagnetic field on production rates of *in situ* produced cosmogenic nuclides. *Earth Planet. Sci. Lett.* **193,** 197–212 (2001).
35. Carcaillet, J. T., Bourles, D. L. & Thouveny, N. Geomagnetic dipole moment and ^{10}Be production rate intercalibration from authigenic ^{10}Be/^{9}Be for the last 1.3 Ma. *Geochem. Geophys. Geosyst.* **5,** doi:10.1029/2003GC000641 (2004).
36. Lejeune, Y. *et al.* Melting of snow cover in a tropical mountain environment: processes and melting. *J. Hydrometeorol.* (in the press).
37. Laskar, J. *et al.* A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428,** 261–285 (2004).
38. Paillard, D., Labeyrie, L. & Yiou, F. Macintosh program performs time-series analysis. *Eos* **77,** 379 (1996).
39. NOAA. *Radiosonde Database Access* ⟨http://raob.fsl.noaa.gov/⟩ (2007).
40. Johannesson, T., Sigurdsson, O., Laumann, T. & Kennett, M. Degree-day glacier mass-balance modeling with applications to glaciers in Iceland, Norway And Greenland. *J. Glaciol.* **41,** 345–358 (1995).
41. Schuler, T. V. *et al.* Distributed mass-balance and climate sensitivity modelling of Engabreen, Norway. *Ann. Glaciol.* **42,** 395–401 (2005).

©2007 Nature Publishing Group

LETTERS

An epipodite-bearing crown-group crustacean from the Lower Cambrian

Xi-guang Zhang[1], David J. Siveter[2], Dieter Waloszek[3] & Andreas Maas[3]

Crown-group crustaceans (Eucrustacea) are common in the fossil record of the past 500 million years back to the early Ordovician period, and very rare representatives are also known from the late Middle and Late Cambrian periods[1]. Finds in Lower Cambrian rocks of the Phosphatocopina, the fossil sister group to eucrustaceans[2], imply that members of the eucrustacean stem lineage co-occurred, but it remained unclear whether crown-group members were also present at that time. 'Orsten'-type fossils are typically tiny embryos and cuticle-bearing animals, of which the cuticle is phosphatized and the material is three-dimensional and complete with soft parts. Such fossils are found predominantly in the Cambrian and Ordovician and provide detailed morphological and phylogenetic information on the early evolution of metazoans. Here we report an Orsten-type Konservat-Lagerstätte from the Lower Cambrian of China that contains at least three new arthropod species, of which we describe the most abundant form on the basis of exceptionally well preserved material of several growth stages. The limb morphology and other details of this new species are markedly similar to those of living cephalocarids, branchiopods and copepods and it is assigned to the Eucrustacea, thus representing the first undoubted crown-group crustacean from the early Cambrian. Its stratigraphical position provides substantial support to the proposition that the main cladogenic event that gave rise to the Arthropoda was before the Cambrian[3]. Small leaf-shaped structures on the outer limb base of the new species provide evidence on the long-debated issue of the origin of epipodites[4,5]: they occur in a set of three, derive from setae and are a ground-pattern feature of Eucrustacea.

Arthropoda
Crustacea
Eucrustacea
Yicaris dianensis gen. et sp. nov.

Etymology. Yi, ethnic minority group of Yunnan Province, China; *caris*, Latin, shrimp; Dian, an ancient kingdom of southern China.

Holotype. Key Laboratory for Palaeobiology, Yunnan University (YKLP), collection number 10840 (Fig. 1f). A trunk end with six appendage-bearing segments and a caudal end with broken furcal rami; an advanced larval stage. The estimated total length of the whole animal is about 1,800 μm.

Referred material. Thirty-four fragmentary topotype specimens (YKLP 10841–10874).

Locality and stratigraphy. Xiaotan section, Yongshan, Yunnan Province; Yu'anshan Formation, *Eoredlichia–Wutingaspis* Biozone (approximately late Atdabanian Stage[6], upper Lower Cambrian). Age determined on diagnostic redlichiid (*Eoredlichia* sp.[7]) and bradoriid (*Kunmingella typica*[8]) arthropods.

Diagnosis. (On the basis of an advanced larval stage.) Eucrustacean with trunk end bearing two laterocaudal lobes, each with up to 15 fine marginal spines: the posterior lobe is on the proximal article of a two-part furca; the distal part of the furca is elongate, paddle-shaped, with about seven setae arising from the gently curved inner margin. Endopod of second antenna to posterior-most developed trunk limb, inclusive, have an inverted cone-shaped distal podomere that has an obliquely faceted distal surface with a palisade of setae.

Description. The head comprises a simple head shield, the ocular and five limb-bearing segments (Figs 1b and 2). Ventrally on the head, the labrum arises in front of the mouth as a bulge-like structure with a steeply raised posterior part (Fig. 1d). The mandibular sternite bears a pair of paragnath humps. Anteriorly, a median lobe separates large paired eye lobes. The first antennae consist of about five podomeres and have a terminal tuft of setae. The second antennae and mandibles both consist of a coxa, a large basipod, an elongate endopod and a multi-annulated exopod (Fig. 1b). The mandibular coxa extends into a hand-shaped median gnathobase. The post-mandibular limbs are biramous; they lack a coxa, but have a large, long basipod that is C-shaped in cross-section (concave posteriorly) and that rests on a finely folded membrane socket (Fig. 1h). The inner margin of the basipod has a seta-bearing proximal endite and a set of up to four or five (in the first and second maxillae) to eight (in post-maxillary limbs) smaller and distally progressively more separated endites, each bearing a row of anterior setae, a short median spine and a crescentic posterior row of setae (Fig. 1k). The exopods of all post-mandibular limbs are leaf-shaped and bear a marginal row of setae (Fig. 1i, j). Up to three lobe-like projections occur on the basipod of each post-mandibular limb except the first maxilla, which has two setae in a similar position. Ontogenetically, each such structure develops from a single seta that widens into a bulb with a vestigial terminal spine and then into a subtriangular to subquadratic flap (Fig. 1f–h) representing an epipodite. The trunk tergites lack lateral extensions but continue ventrally into the finely folded softer body proper (Figs 1h and 2).

Six consecutive growth stages are represented, having 4–9 trunk tergites, respectively. The earliest known stage is approximately 460 μm long (Fig. 1a). The tail portion is not separated from the last tergite-bearing segment and continues into a pair of furcal rami. Only the first trunk limb is developed; the three more posterior limbs are progressively less developed; the most-posterior limb is an almost laterally inserted, bifid-lobe-shaped bud.

Trunk segments and limbs are progressively added from stage to stage and limb endites get larger and develop more setal armature (Fig. 1a–e). Fragmentary material indicates that a mature limb was at least twice as long as that of a growth stage with 8 thoracomeres and had more than 200 setae on its median enditic surface. Our sixth growth stage is about 1.8 mm long; the total length of the adults is

[1]Key Laboratory for Palaeobiology, Yunnan University, Kunming 650091, China. [2]Department of Geology, University of Leicester, Leicester LE1 7RH, UK. [3]Section for Biosystematic Documentation, University of Ulm, D-98081 Ulm, Germany.

©2007 Nature Publishing Group

estimated to be more than 3 mm. In late-stage larvae the furca is articulated and two-segmented. The head shield and eyes show no major ontogenetic changes.

Discussion. The material of *Y. dianensis* is important in two respects. First, it displays the post-embryonic ontogeny of an animal as old as the Early Cambrian. Second, *Y. dianensis* is temporally close to evolutionary events deep in arthropod and, specifically, crustacean phylogeny. Its development should, therefore, be closer to the original developmental pattern of the stem forms and less changed than that of recent in-group taxa, which have accumulated lineage-specific modifications. With such ontogeny data and having the age of the fossil as a time marker, it is possible to more precisely include ontogenetic evolutionary pathways in the reconstruction of relationships and ground patterns of stem species and monophyla. According to our analysis, *Y. dianensis* represents the first undoubted eucrustacean known from the Lower Cambrian (the single previously described Lower Cambrian putative eucrustacean species[9] lacks eucrustacean characters[1]) and can serve as a substantial tool for testing relevant character acquisition and phylogenetic hypotheses. This is of particular importance because crustacean phylogeny has gained new interest by recent studies using neurobiological, developmental-biological and molecular investigations[10–15].

Yicaris dianensis shows greatest similarities with the upper Middle Cambrian eucrustacean branchiopod *Rehbachiella kinnekullensis*[16–18] and the living minute and blind cephalocarids, both in its head and trunk-limb design. *Y. dianensis* differs from *R. kinnekullensis* in aspects of larval development (for example, retaining prominent eye lobes throughout ontogeny), in having a head-restricted shield

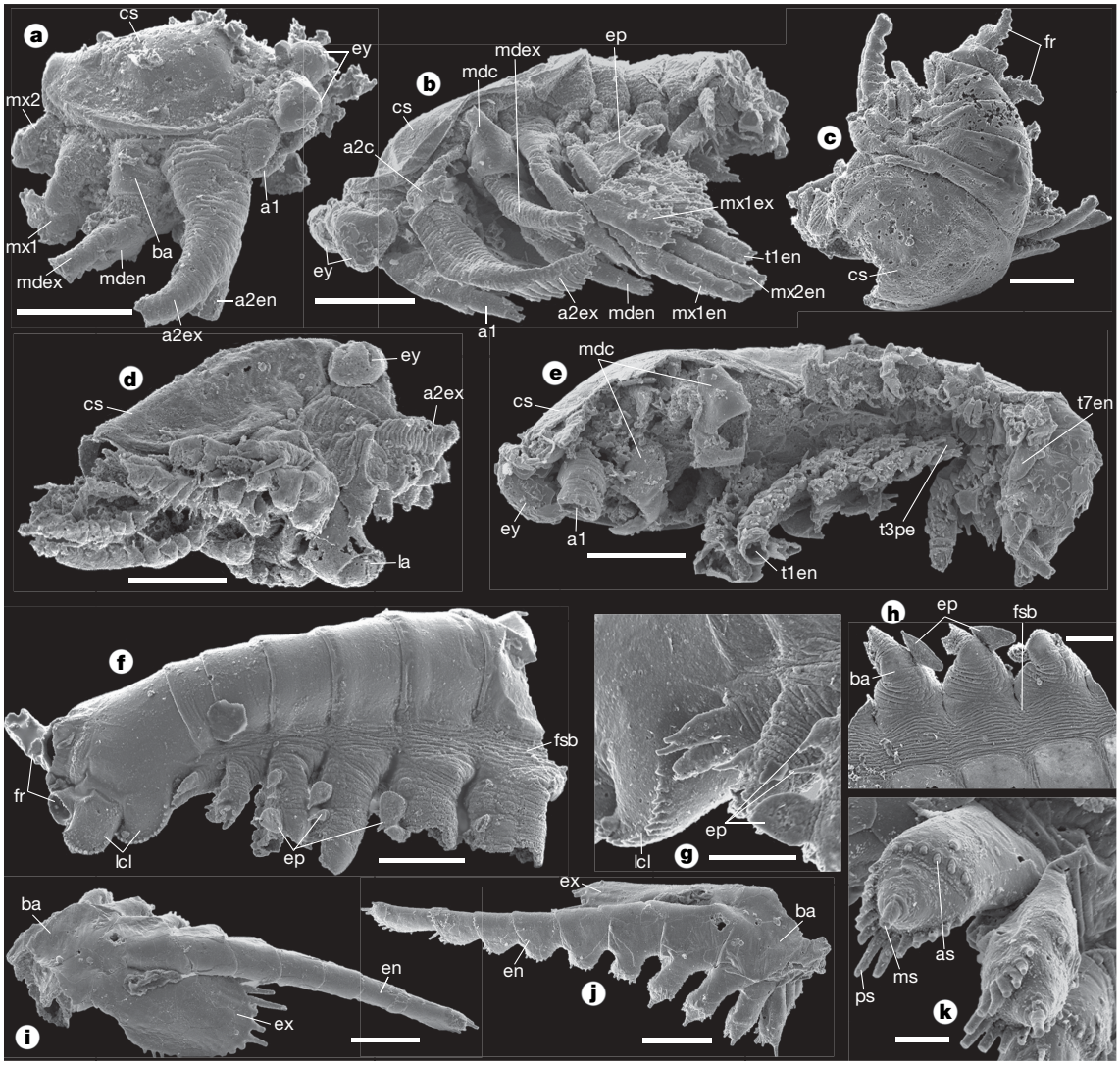


Figure 1 | *Y. dianensis*; topotype specimens. **a**, Small larva, only head portion preserved (developmental stage 1); oblique lateral view (YKLP 10841). **b**, Reasonably complete, distorted specimen (developmental stage 2); lateral view (YKLP 10844). **c**, Distorted larva with five trunk segments (developmental stage 2); oblique anterolateral view (YKLP 10842). **d**, Distorted larva, only head portion preserved (developmental stage 2); oblique ventral view (YKLP 10843). **e**, Reasonably complete distorted specimen (developmental stage 3); oblique ventral view (YKLP 10845). **f**, **g**, Fragment of trunk (holotype; developmental stage 6); lateral view (**f**) and detail of epipodites on basipods (**g**) (YKLP 10840). **h**, Fragment of right side of trunk, comprising four-plus thoracomeres (developmental stage uncertain); ventral view (YKLP 10847). **i**–**k**, Single post-maxillulary limb, proximal-most part with several endites broken off (developmental stage 6); anterolateral (**i**) and anterior (**j**) views, and two endites of basipod (**k**) (YKLP 10859). All scale bars represent 100 μm, except **g** (50 μm) and **k** (20 μm). a1, first antenna; a2c, coxa of second antenna; a2en, endopod of second antenna; a2ex, exopod of second antenna; as, anterior enditic setae; ba, basipod; cs, cephalic shield; en, endopod; ep, epipodite(s); ex, exopod; ey, eye(s); fr, furcal rami; fsb, folded soft body; la, labrum; lcl, laterocaudal lobe(s); mdc, mandibular coxa; mden, endopod of mandible; mdex, exopod of mandible; ms, median enditic seta or spine; mx1, first maxilla; mx1en, endopod of first maxilla; mx1ex, exopod of first maxilla; mx2, second maxilla; mx2en, endopod of second maxilla; ps, posterior enditic setae; t1en, endopod of first trunk limb; t3pe, proximal endite of third trunk limb; t7en, endopod of seventh trunk limb.

©2007 Nature Publishing Group

and lacking a filter groove between the trunk limbs (a speciality of branchiopods). It differs from the cephalocarids in having eyes and lacking tergopleural outgrowths in the thorax. Like early larvae of cephalocarids, maxillopods and especially branchiopods, *Y. dianensis* has a prominent second antenna with a large seta-bearing exopod for locomotion and for sweeping in food and has spine-bearing endites surrounding the labrum. In all larval and adult malacostracans, the second antenna lacks any structures for food intake. Another similarity with fossil and living branchiopods[16,17] and cephalocarids[19] is the form of the post-mandibular limbs with their elongate, rather fleshy basipods armed with up to seven setiferous, soft endites medially. Such endites are unknown from the post-mandibular limbs of malacostracans, and they are also absent in stem crustaceans and in the Phosphatocopina[20,21]. Furthermore, the basipod of the first maxilla of *Y. dianensis* has four endites, like that in living maxillopods, larval cephalocarids and *Rehbachiella*; however, in the first maxilla of malacostracans, both coxa and basipod are drawn out into a single blade-like endite. The basipods of phosphatocopines and stem crustaceans have only one to two median, stiff enditic prolongations[20].

As indicated by its prominent second antenna with its large multi-annulated exopod, the leaf-shaped post-mandibular exopods armed with long marginal setae, and the setiferous furcal rami, *Y. dianensis* was probably capable of swimming. Its post-mandibular limb apparatus possibly functioned in a metachronal beat for both swimming and food intake. This is the case in cephalocarids[19] and branchiopods[16,22], in which the limb apparatus forms an open-chamber system with food approaching the postoral mandibular coxae from behind along the body surface. Another detailed similarity between *Y. dianensis*, cephalocarids and branchiopods within this apparatus is the presence of the C-curved basipods that form chambers that are opened and closed with the metachronal beat, thus facilitating a sucking system for food gathering and an outwardly directed water current for locomotion[23], directed by the exopods (and epipodites if developed). In all these taxa, food is passed forward by the proximal endites. Limb designs and food-gathering systems are completely different in malacostracans[16,24–26], reflecting different methods of food intake. Other characters associated with feeding, such as a labrum, a coxa on the second antenna and mandible, paragnath humps and the tripartite enditic setal armature[21], occur in *Y. dianensis*, phosphatocopines and eucrustaceans (and, in part, possibly even in myriapods and insects[14]), but are lacking in stem crustaceans. Phosphatocopines lack the specialized first maxilla known from eucrustaceans[2]. With regards to phylogeny, the most parsimonious position is that *Yicaris* cannot be placed below phosphatocopines. As an eucrustacean, its morphology supports a monophyletic origin of cephalocarids, branchiopods and maxillopods (the traditional Entomostraca[1]), possibly forming the sister group of Malacostraca (Fig. 3). Inclusion of insects or myriapods does not contradict this relationship hypothesis, possibly as the sister group to eucrustaceans.

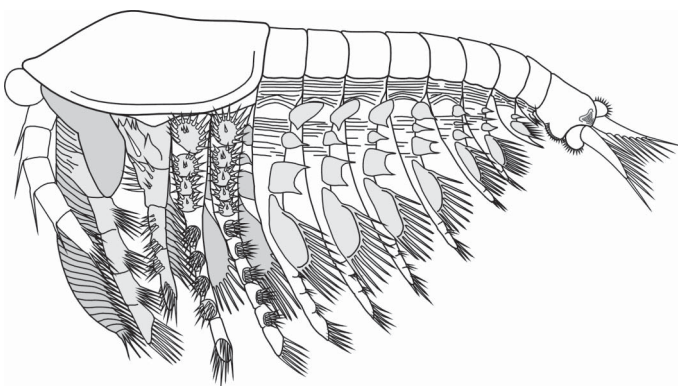


Figure 2 | Reconstruction of *Y. dianensis* on the basis of a larval stage with eight thoracomeres (developmental stage 4). The head limbs are drawn in median view and the trunk limbs in lateral view.

The morphology of *Yicaris* also provides fundamental insight regarding the enigmatic developmental origin of epipodites[4,5,16], indicating that Eucrustacea basally possessed a set of three epipodites. These soft, thin, lobe- to leaf-shaped osmoregulatory and/or respiratory outgrowths occur on the outer side of limb bases of various eucrustaceans. They are unknown, for example, in remipedes[27], cirripedes, mystacocarids and all hitherto described Cambrian eucrustaceans, probably being secondarily lost. In *Yicaris*, the epipodites develop from spine-like setae that ontogenetically extend proximally to become bulbous and eventually leaf-shaped with a marginal setal remnant. This evidence may also help to clarify that the outer proximal 'leaf' on the exopod of cephalocarids, variously regarded as 'epipodite', the 'pseudo-epipodite' or part of the exopod[16,19], is possibly the remnant of the distal-most of three epipodites. Living branchiopods have retained up to two epipodites, which develop from small seta-less 'lobes'[17]. The epipodites of leptostracan and stomatopod malacostracans[28] can be likewise homologized. Epipodites and a sophisticated combined locomotory and food-gathering apparatus, as present in *Yicaris* and, presumably, in the eucrustacean ancestor, may have been a significant factor leading to the successful diversification of eucrustaceans already by the Cambrian. Some authors consider that insect wings may have originated from epipodites[29,30]; the early Cambrian occurrence of this trait is of potential significance to the debate regarding the emergence of winged (pterygote) forms within euarthropods.

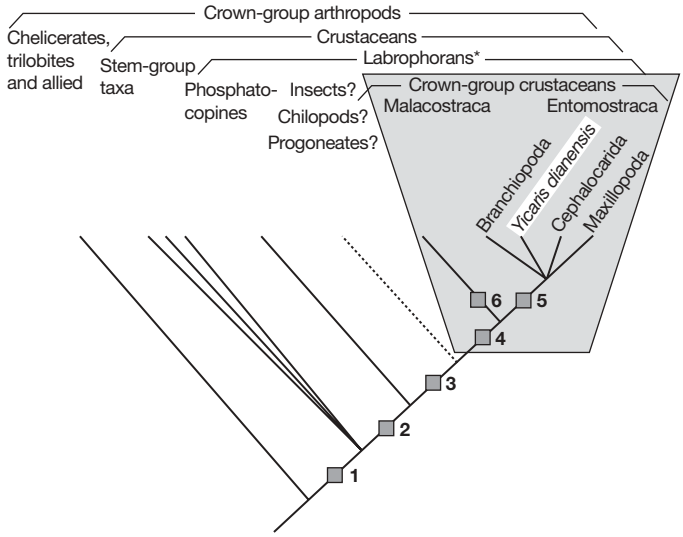


Figure 3 | Proposed phylogenetic position of *Y. dianensis*, within the crown group of arthropods. Numbers refer to examples of autapomorphies at major nodes. 1, Crustacea: basipod of post-antennular limbs with a movable, seta-bearing endite (the proximal endite); second and third head limbs modified in having a multi-annulated exopod with inward-pointing setae. 2, Labrophora: five limb-bearing head segments; labrum; paragnaths; proximal endite of second limb and third limb (mandible) enlarged to each form a coxa (asterisk: if including the ateloceratans, Labrophora would be equivalent to Mandibulata). 3, (unnamed): first post-mandibular limb specialized as a mouthpart, the first maxilla (= maxillula). 4, Eucrustacea: first larval stage a nauplius with only three limb-bearing segments. 5, Entomostraca: basipods of post-mandibular limbs long and fleshy, drawn out into a set of soft setae-bearing endites medially; distal part of mandible (basipod plus rami) lost during late ontogeny; basal part of first maxilla (maxillula) with four soft endites medially (proximal endite plus three on the basipod). 6, Malacostraca: numerous autapomorphies, such as body tagmosis (limb-bearing trunk of two parts, with eight and six segments, respectively); genital opening on trunk segment six of females and on trunk segment eight of males; first antenna with two flagellae, second antenna with endopod modified into flagellum; distal part of mandible modified into a tripartite structure; first and second maxillae flattened, serving as mouthparts. For further justification of the phylogram, see Supplementary Information.

©2007 Nature Publishing Group

METHODS SUMMARY

All material was recovered from residues obtained by 5% acetic acid digestion of carbonate nodules from black shales.

Received 16 April; accepted 30 July 2007.

1. Waloszek, D. in *The New Panorama of Animal Evolution* (eds Legakis, A., Sfenthourakis, S., Polymeni, R. & Thessalou-Legaki, M.) 69–87 (Pensoft, Sofia-Moscow, 2003).
2. Siveter, D. J., Waloszek, D. & Williams, M. An early Cambrian phosphatocopid crustacean with three-dimensionally preserved soft parts from Shropshire, England. *Spec. Pap. Palaeont.* **70,** 9–30 (2003).
3. Peterson, K. J., McPeek, M. A. & Evans, D. A. Tempo and mode of early animal evolution: inferences from rocks, Hox, and molecular clocks. *Paleobiology* **31** (suppl.), 36–55 (2005).
4. Lauterbach, K. Über die mutmaßliche Herkunft der Epipodite der Crustacea. *Zool. Anz.* **202,** 33–50 (1979).
5. Hessler, R. R. Reflections on the phylogenetic position of the Cephalocarida. *Acta Zool.* **73,** 315–316 (1992).
6. Hou, X.-G. *et al. The Cambrian Fossils of Chengjiang, China: the Flowering of Early Animal Life* (Blackwell, Oxford, 2004).
7. Luo, H.-L. *et al. The Sinian–Cambrian Boundary in Eastern Yunnan, China* (Yunnan People's Publishing, Kunming, 1982).
8. Hou, X.-G., Siveter, D. J., Williams, M. & Feng, X.-H. A monograph of Bradoriid arthropods from the Lower Cambrian of southwest China. *Trans. R. Soc. Edinb. Earth Sci.* **92,** 347–409 (2002).
9. Chen, J.-Y., Vannier, J. & Huang, D.-Y. The origin of crustaceans: new evidence from the Early Cambrian of China. *Proc. R. Soc. Lond. B* **268,** 2181–2187 (2001).
10. Babbitt, C. C. & Patel, N. H. in *Crustacea and Arthropod Relationships* Vol. 16 (eds Koenemann, S. & Jenner, R. A.) 275–294 (Taylor & Francis, Boca Raton, 2005).
11. Fanenbruck, M. & Harzsch, S. A brain atlas of *Godzilliognomus frondosus* Yager, 1989 (Remipedia, Godzilliidae) and comparison with the brain of *Speleonectes tulumensis* Yager, 1987 (Remipedia, Speleonectidae): implications for arthropod relationships. *Arth. Struc. Dev.* **34,** 343–378 (2005).
12. Giribet, G., Richter, S., Edgecombe, G. D. & Wheeler, W. C. in *Crustacea and Arthropod Relationships* Vol. 16 (eds Koenemann, S. & Jenner, R. A.) 307–352 (Taylor & Francis, Boca Raton, 2005).
13. Glenner, H., Thomsen, P. F., Hebsgaard, M. B., Sørensen, M. V. & Willerslev, E. The origin of insects. *Science* **314,** 1883–1884 (2006).
14. Wolff, C., &. Scholtz, G. Cell lineage analysis of the mandibular segment of the amphipod *Orchestia cavimana* reveals that the crustacean paragnaths are sternal outgrowths and not limbs. *Front. Zool.* **3,** 1–14 (2006).
15. Carapelli, A. *et al.* in *Crustacea and Arthropod Relationships* Vol. 16 (eds Koenemann, S. & Jenner, R. A.) 295–306 (Taylor & Francis, Boca Raton, 2005).
16. Walossek, D. The Upper Cambrian *Rehbachiella* and the phylogeny of Branchiopoda and Crustacea. *Foss. Strat.* **32,** 1–202 (1993).
17. Olesen, J. in *Evolutionary Developmental Biology of Crustacea* Vol. 15 (ed. Scholtz, G.) 217–269 (Lisse, Balkema, 2004).
18. Schram, F. R. & Koenemann, S. Developmental genetics and arthropod evolution: part I, on legs. *Evol. Dev.* **3,** 343–354 (2001).
19. Sanders, H. L. The Cephalocarida. Functional morphology, larval development, comparative external anatomy. *Mem. Conn. Acad. Arts Sci.* **15,** 1–80 (1963).
20. Maas, A., Waloszek, D. & Müller, K. J. Morphology, ontogeny and phylogeny of the Phosphatocopina (Crustacea) from the Upper Cambrian "Orsten" of Sweden. *Foss. Strat.* **49,** 1–238 (2003).
21. Walossek, D. & Müller, K. J. Stem-lineage crustaceans from the Upper Cambrian of Sweden and their bearing upon the position of *Agnostus. Lethaia* **23,** 409–427 (1990).
22. Cannon, H. G. & Leak, F. M. C. On the feeding mechanism of the Branchiopoda. *Phil. Trans. R. Soc. Lond. B* **222,** 267–352 (1933).
23. Barlow, D. I. & Sleigh, M. A. in *The Brine Shrimp Artemia, 1, Morphology, Genetics, Radiobiology, Toxicology* (eds Persoone, G., Sorgeloos, P., Roels, O. & Jaspers, E.) 61–73 (Universa, Wetteren, 1980).
24. Cannon, H. G. On the feeding mechanism of *Nebalia bipes. Trans. R. Soc. Edinb.* **55,** 355–369 (1927).
25. Cannon, H. G. Nebaliacea. *Discov. Rep.* **3,** 199–222 (1931).
26. Hamner, W. M. Biomechanics of filter feeding in the Antarctic krill *Euphausia superba*: review of past work and new observations. *J. Crust. Biol.* **8,** 149–163 (1988).
27. Yager, J. Remipedia, a new class of Crustacea from a marine cave in the Bahamas. *J. Crust. Biol.* **1,** 328–333 (1981).
28. Burnett, B. R. & Hessler, R. R. Thoracic epipodites in the Stomatopoda (Crustacea): a phylogenetic consideration. *J. Zool.* **169,** 381–392 (1973).
29. Averof, M. & Cohen, S. M. Evolutionary origin of insect wings from ancestral gills. *Nature* **385,** 627–630 (1997).
30. Boxshall, G. A. The evolution of arthropod limbs. *Biol. Rev.* **79,** 253–300 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Natural Science Foundation of China, the Ministry of Science and Technology of China, the Royal Society and the German Research Council.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to X.-g.Z. (xgzhang@ynu.edu.cn).

©2007 Nature Publishing Group

LETTERS

Reconciling complexity with stability in naturally assembling food webs

Anje-Margriet Neutel[1]†, Johan A. P. Heesterbeek[2], Johan van de Koppel[3], Guido Hoenderboom[4], An Vos[5], Coen Kaldeway[5], Frank Berendse[6] & Peter C. de Ruiter[5,7]

Understanding how complex food webs assemble through time is fundamental both for ecological theory and for the development of sustainable strategies of ecosystem conservation and restoration. The build-up of complexity in communities is theoretically difficult, because in random-pattern models complexity leads to instability[1]. There is growing evidence, however, that nonrandom patterns in the strengths of the interactions between predators and prey strongly enhance system stability[2–4]. Here we show how such patterns explain stability in naturally assembling communities. We present two series of below-ground food webs along natural productivity gradients in vegetation successions[5,6]. The complexity of the food webs increased along the gradients. The stability of the food webs was captured by measuring the weight of feedback loops[7] of three interacting 'species' locked in omnivory. Low predator–prey biomass ratios in these omnivorous loops were shown to have a crucial role in preserving stability as productivity and complexity increased during succession. Our results show the build-up of food-web complexity in natural productivity gradients and pin down the feedback loops that govern the stability of whole webs. They show that it is the heaviest three-link feedback loop in a network of predator–prey effects that limits its stability. Because the weight of these feedback loops is kept relatively low by the biomass build-up in the successional process, complexity does not lead to instability.

Ecologists have long studied processes of community assembly and have revealed general principles governing community diversity, dynamics and functioning in vegetation successions[8], that is, on a single trophic level. The understanding of communities at multi-trophic levels, however, is much less advanced. In processes of ecological succession, increasing productivity could give longer food chains[9–11], although opposite trends have also been found[12], and it has been argued that stability constraints may limit food-chain length[13]. Moreover, the assembly of complex food webs is not self-evident, because random community models suggest that complexity promotes instability[1]. However, there has been increasing evidence that the nonrandom patterning of strong and weak links in food webs greatly enhances the stability of these networks[2–4]. It has been suggested that such stabilizing patterns can be caused by the decrease in biomass over increasing trophic levels, which means that long feedback loops in the food webs contain relatively weak links[7]. In a comparison between steady states of food-web models varying in complexity, and with hypothetical biomass pyramids, it was shown that complexity does not enhance instability[7].

So far, there have been no studies that show and quantify the build-up of complexity of food webs in long-term successional processes. The question still stands: do food webs become more complex as ecosystems develop and, if so, how can these systems exist—what keeps them stable, despite their increasing complexity?

Here we present a series of food webs sampled over two successional gradients along which natural productivity varies. We show that the food webs increase in complexity and that it is not their complexity as such, but the organization of link strengths over feedback loops, measured as 'maximum loop weight' of omnivorous loops[7], that governs their stability. We do this within the classical framework of May[1]. Maximum loop weight has been shown to characterize and explain the natural organization of food webs by distinguishing empirically patterned webs from their randomized counterparts[7]. However, so far the approach has not been used to quantify differences between various empirical food webs, nor to study the process of community assembly. This is what we will do here.

We studied two successions of below-ground food webs in sandy dune soils: one on the Waddensea Island of Schiermonnikoog in the north of the Netherlands[5] and the other at Hulshorsterzand, on the Veluwe, in the central Netherlands[6]. The study sites, which constitute the two gradients, represent four consecutive stages in chronosequences of early primary vegetation succession, increasing in above-ground[5,6] and below-ground productivity (see Methods). The observed soil organisms were categorized as 'functional groups' on the basis of taxonomy, life-history traits and diet[14,15]. We regarded observed yearly average biomass densities (see Methods and Supplementary Table 1) as steady states, that is, as stage-equilibria interpreted in the sense of ref. 9.

We found an increase in complexity of the food webs along both gradients. Food-web diagrams representing the four successional stages illustrate this increase (Fig. 1). When we quantify this, we see that the number of trophic groups (n), the maximum food-chain length and the link density (nC, where C is connectance[1,16]) all increased along the gradients (Supplementary Table 2). There was not a strict growth of the webs, because in the last stage phytophagous nematodes tended to disappear (Fig. 1).

We calculated the interaction strengths—effects of the trophic groups on each other—using the methods from refs. 2 and 17 (see Methods). We determined food-web stability (that is, the ability to return to the steady state after a small perturbation) using the method from ref. 7 (see Methods).

We examined how food-web structure is related to system stability. To do this, we first related stability to two previously proposed measures of food-web complexity that average over all

[1]Environment Department, University of York, Heslington, York YO10 5DD, UK. [2]Faculty of Veterinary Medicine, Theoretical Epidemiology, Utrecht University, 3508 TD Utrecht, The Netherlands. [3]Spatial Ecology Department, Centre for Estuarine and Marine Ecology, Netherlands Institute of Ecology (NIOO-KNAW), 4400 AC Yerseke, The Netherlands. [4]Department ICT, Wageningen University and Research Centre, 6700 AB Wageningen, The Netherlands. [5]Alterra, Soil Science Centre, Wageningen University and Research Centre, 6700 AA Wageningen, The Netherlands. [6]Nature Conservation and Plant Ecology Group, Wageningen University, 6708 PB Wageningen, The Netherlands. [7]Department of Environmental Sciences, Copernicus Institute for Sustainable Development and Innovation, Utrecht University, 3508 TC Utrecht, The Netherlands. †Present address: British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

©2007 Nature Publishing Group

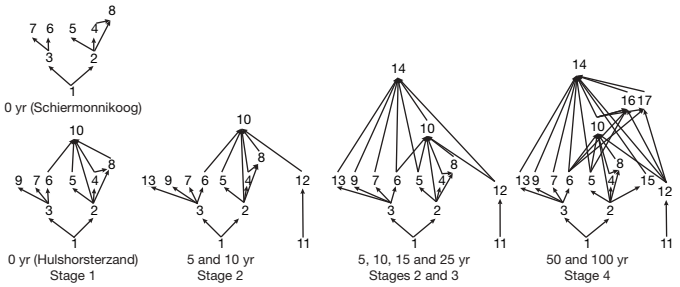


Figure 1 | Observed connectedness structures of the below-ground food webs of Schiermonnikoog and Hulshorsterzand, representative of the four successional stages. The numbers refer to the trophic groups: 1, detritus; 2, bacteria; 3, fungi; 4, flagellates; 5, bacterivorous nematodes; 6, fungivorous nematodes; 7, noncryptostigmatic mites; 8, amoebae; 9, cryptostigmatic mites; 10, predatory nematodes; 11, roots; 12, phytophagous nematodes; 13, collembolans; 14, predatory mites; 15, bacterivorous mites; 16, nematophagous mites; and 17, predatory collembolans. Years refer to soil age (see Methods). For details and exceptions, see Supplementary Table 1.

predator–prey interactions: a measure of link density (nC) and an analogue of May's complexity measure[1] $a(nC)^{\frac{1}{2}}$, where a is the average interaction strength (see Methods). Then we related stability to a measure of food-web structure that incorporates the organization of interaction strengths: the maximum weight (see Methods) of omnivorous loops[7]. The measure does not include the shortest feedback loops, the predator–prey loops of two links (generating negative feedback), because they do not contribute to instability in these systems[7,18].

We found no strong and consistent relationship between link density nC and stability in the two series (Fig. 2a). Incorporating the average strengths of the interactions did not explain the stability of the food webs either; there was also no consistent relationship between the analogue of May's complexity measure $a(nC)^{\frac{1}{2}}$ and stability (Fig. 2b).

We calculated the weights of the short two-link feedback loops and of omnivorous loops in the food webs (see Methods). In all the webs, loop weights decreased with loop lengths; the maximum weight of omnivorous loops in all food webs was that of loops with a length of three, with two top-down effects and one bottom-up effect. This was in correspondence with the results reported in ref. 7. This is a positive feedback loop, because it consists of two negative effects and one positive effect.

We found that the maximum weight of omnivorous loops was strongly related to food-web stability in both series (Fig. 2c). A sensitivity analysis carried out with respect to observed biomass densities and to our way of deriving stability revealed the robustness of these results (see Supplementary Figs 1, 2 and 3).

The observed relationship between stability and the maximum weight of omnivorous loops does not depend on the specific food-web types and configurations or on the range of complexity observed in these webs. This was illustrated by modelling the large set of food webs assembled in ref. 19 (Fig. 3; for parameterization, see Methods). We chose this set—acknowledging that it does not provide the most detailed data on real webs available—not to assess the actual stability of these food webs, but to test our relationship in a large series of webs from different environments, and with a wide variety of architectures.

On the basis of these results, we propose that stability in predator–prey communities (with at least one three-link loop) is limited by the maximum weight of (positive feedback) loops of length three, in biological terms described by:

$$\max_{E_3}\left\{\left(\frac{f_{23}}{d_3}\frac{f_{12}}{d_2}e_{13}\frac{f_{13}}{d_1}\frac{B_3}{B_1}\right)^{\frac{1}{3}}\right\} \quad (1)$$

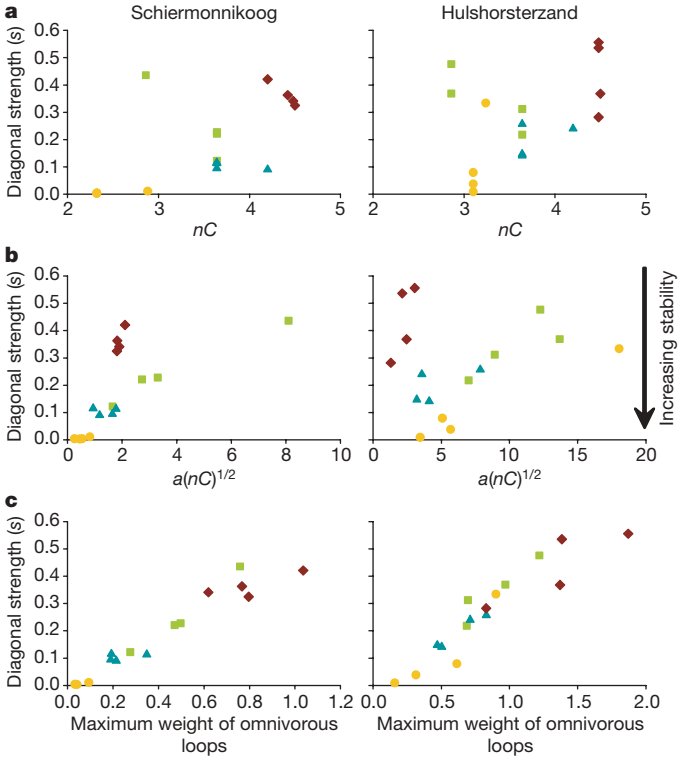


Figure 2 | Food-web stability related to measures of food-web structure, of Schiermonnikoog and Hulshorsterzand. **a**, Link density nC; **b**, $a(nC)^{\frac{1}{2}}$; **c**, maximum weight of omnivorous loops. The successional stages (1, yellow circles; 2, green squares; 3, blue triangles; and 4, brown diamonds) were each represented by four food-web replications. All quantities are dimensionless. For details, see Methods. Linear regression: **a**, $y = -0.24 + 0.12x$, $r^2 = 0.34$, $P = 0.018$ (Schiermonnikoog) and $y = -0.16 + 0.12x$, $r^2 = 0.19$, $P = 0.096$ (Hulshorsterzand); **b**, $y = 0.07 + 0.06x$, $r^2 = 0.44$, $P = 0.005$ (Schiermonnikoog) and $y = 0.23 + 0.006x$, $r^2 = 0.03$, $P = 0.497$ (Hulshorsterzand); and **c**, $y = -0.007 + 0.47x$, $r^2 = 0.94$, $P < 0.001$ (Schiermonnikoog) and $y = -0.03 + 0.35x$, $r^2 = 0.87$, $P < 0.001$ (Hulshorsterzand). The sample size is $N = 16$ in all statistical analyses.

where E_3 is the set of omnivorous loops of length three; the subscripts 1, 2, and 3 refer to bottom prey, intermediate predator and top predator in the loop, respectively; f_{ij} is a mass-specific predation rate

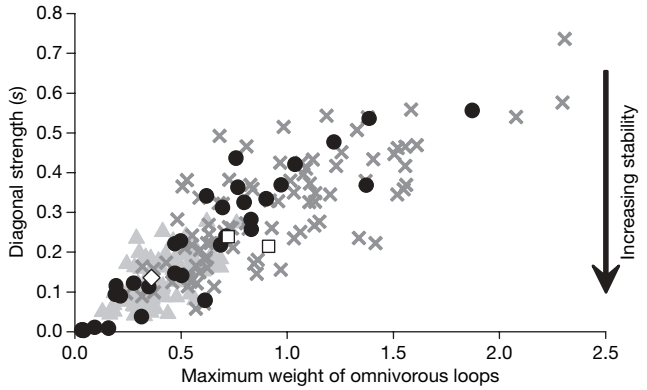


Figure 3 | Relationship between loop weight and stability for 'Cohen webs' compared with observed webs (black circles). Cohen webs were modelled with 10-fold (grey crosses) and 20-fold (grey triangles) biomass decrease per trophic level (Methods). In 187 of 190 webs studied, the loop with maximum weight had length 3 (with 2 top-down effects and 1 bottom-up effect); in 3 it had length 4 (2 with 10-fold (white squares) and 1 with 20-fold biomass decrease (white diamond)). nC ranged from 2.3 to 8.8, and n ranged from 14 to 99. Linear regressions: $y = 0.03 + 0.28x$, $r^2 = 0.71$, $N = 190$, $P < 0.001$ (Cohen webs); $y = 0.01 + 0.34x$, $r^2 = 0.85$, $N = 32$, $P < 0.001$ (observed webs).

©2007 Nature Publishing Group

of predator j on its prey i in equilibrium; e_{ij} is the efficiency with which j converts food into biomass, from feeding on i, $0 < e_{ij} < 1$; d_i is the total natural (that is, non-predatory) specific death rate of i in equilibrium (a scaling factor, used for consistency with the stability metric, which is also relative to these death rates, see Methods); and B_j is the equilibrium biomass of predator j.

Expression (1) gives a clear insight into how organism traits and biomass structure affect loop weight and, hence, stability. Low top–bottom biomass ratios ($\frac{B_2}{B_1}$) contribute to low loop weight, as does skewed feeding of the top predator on intermediate and bottom prey ($f_{23}f_{13}$). The latter effect could, for example, be enhanced by low biomass ratios of intermediate and bottom prey if the omnivores feed according to prey biomass.

We identified the predator–prey biomass ratios on which expression (1) directly depended for all the observed food webs (Supplementary Information). We found that the lower these predator–prey ratios were, the more stable the webs (Fig. 4). Moreover, in 21 of the 32 webs it was the same loop that had maximum (omnivorous) weight: that of predatory nematodes, bacterivorous nematodes and bacteria (crossed circles, Fig. 4). This showed how a single predator–prey biomass ratio had a key role in governing stability along the gradients. It also revealed a pattern over time in both gradients: this biomass ratio increased from the first to the second stage, when predatory nematodes were the top predators, and decreased again with complexity in the later stages, when predatory nematodes were no longer the top predators (Supplementary Fig. 4). In the last stage, this ratio was relatively low (Supplementary Fig. 4), but now loops higher in the web became the heaviest and new biomass ratios determined stability (Fig. 4).

The results suggest an alternating pattern of decreasing and increasing stability in the course of ecosystem succession: biomass is being built up on the top trophic level, causing an increase in maximum loop weight and a decrease in stability. New structure is created by the entrance of a new top predator. This causes predation pressure of the underlying trophic level on lower prey to decrease, which, in turn, decreases maximum loop weight and increases stability. Then, a further build-up of biomass on the highest trophic levels makes a new omnivorous loop the heaviest, increasing maximum loop weight and decreasing stability again. The observed pattern seems to reveal a cascading effect of increasing food-chain length with increasing productivity, as predicted in ref. 10. It suggests that top-down control operates not only within relatively simple food chains, but also within complex networks of intertwined chains[20,21]. Furthermore, this top-down control may have a key role in community assembly: the biomass build-up over succession does not lead to instability because predation is dampened again by other, higher trophic-level predators, thus solving May's paradox.

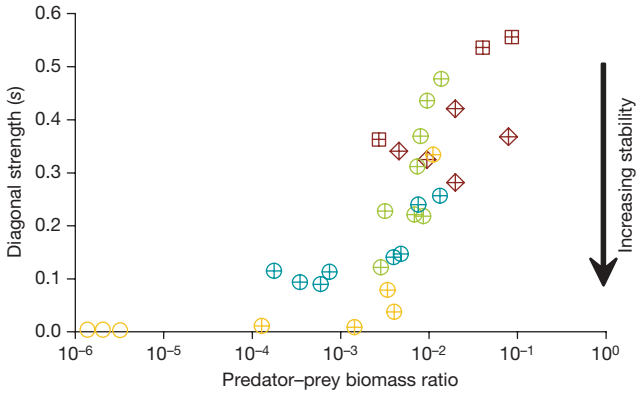


Figure 4 | Relationship between predator–prey biomass ratios and stability in the food webs along both successional gradients (Schiermonnikoog and Hulshorsterzand combined). Predator–prey biomass ratios are defined as the biomass ratios in a food web directly determining expression (1) (that is, determining interaction strengths in the loop with maximum omnivorous weight). These were $\frac{B_{10}}{B_2}$ (crossed circles), $\frac{B_4 B_8}{\left(\sum_{i\in\Delta_8} B_i\right)^2}$ (open circles), $\frac{B_{10}B_{16}}{\left(\sum_{i\in\Delta_{16}} B_i\right)^2}$ (crossed diamonds) and $\frac{B_{10}B_{17}}{\left(\sum_{i\in\Delta_{17}} B_i\right)^2}$ (crossed squares), where Δ_j is the set of indices referring to the prey species of j, and subscripts refer to the trophic group (see Fig. 1). The colours refer to successional stage, as in Fig. 2. Linear regression of the relationship between $\frac{B_{10}}{B_2}$ and stability: $y = 0.05 + 26x$, $r^2 = 0.71$, $N = 21$, $P < 0.001$.

The emergence and persistence of biomass pyramids over succession has a clear basis in observation and theory[10,11,22,23] and should be a general feature of most natural systems. Our results indicate that the shape of the biomass pyramid may have a crucial role in food-web stability. They suggest that more biomass at the top of food chains will limit stability through the weight of three-link (positive) feedback loops. Our simple metric emphasizes the importance of interaction strengths in the smallest omnivorous structures[3,24].

In ecological networks, stability depends on the relationship between the positive and negative feedback loops, on one hand, and the self-damping within populations, on the other. By quantifying the feedback loops, our analysis takes a step in unravelling this relationship between interspecific and intraspecific forces. It is generally known that positive feedback loops are destabilizing and negative feedback loops are stabilizing. Our results suggest that the Achilles heel of a predator–prey network is its heaviest positive feedback loop. Thus, the approach may be of value for a broader class of ecological networks, in which the interactions are known to form skewed patterns of strong and weak links[25,26].

METHODS SUMMARY

We took soil samples in the upper soil layers of each of the four stages in both vegetation successions in four replications. Soil samples were taken three times within a year (resulting in a total of 96 soil samples) to establish yearly average biomass densities.

We used the yearly averages of the biomass densities to calculate the interaction strengths—elements of the (jacobian) community matrices[1] representing the soil food webs—using the methods from refs. 2 and 17. To do this, we regarded observed yearly average biomass densities as steady states—that is, states in which biomass growth rates of each population (gain through feeding) equal death rates (loss through natural death and predation). Effects of predators on their prey (negative interaction strengths) were mass-specific predation rates at steady state, and effects of prey on their predator (positive interaction strengths) were growth rates, per prey biomass, at steady state[2,4,7,13]. We calculated steady-state predation rates by using empirical values of total-mass-specific natural death rates, conversion efficiencies and prey preferences (rough estimations of prey's proportions in a predator's diet, relative to prey abundance).

We determined stability of the soil food webs using the method from ref. 7. Stability was calculated by assessing the level of intraspecific interaction (diagonal strength) needed for all eigenvalues in a community matrix to have negative real parts. This level of intraspecific interaction was expressed as the proportion of total-mass-specific natural (that is, non-predatory) death at the steady state that was density-dependent. Food webs that required less diagonal strength were called 'more stable'.

For the comparison of stability with measures of food-web structure, we calculated average interaction strength and loop weights[7]. Both quantities were scaled for total-mass-specific natural death rates.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 April; accepted 7 August 2007.

1. May, R. M. Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
2. de Ruiter, P. C., Neutel, A. M. & Moore, J. C. Energetics, patterns of interaction strengths, and stability in real ecosystems. *Science* **269**, 1257–1260 (1995).
3. McCann, K., Hastings, A. G. & Huxel, R. Weak trophic interactions and the balance of nature. *Nature* **395**, 794–798 (1998).
4. Rooney, N., McCann, K., Gellner, G. & Moore, J. C. Structural asymmetry and the stability of diverse food webs. *Nature* **442**, 265–269 (2006).
5. Olff, H., Huisman, J. & van Tooren, B. F. Species dynamics and nutrient accumulation during early primary succession in coastal sand dunes. *J. Ecol.* **81**, 693–702 (1993).
6. Berendse, F. & Elberse, W. Th. in *Perspectives on Plant Competition* (eds Grace, J. B. & Tilman, D.) 94–116 (Academic, New York, 1990).

©2007 Nature Publishing Group

7. Neutel, A. M., Heesterbeek, J. A. P. & de Ruiter, P. C. Stability in real food webs: weak links in long loops. *Science* **296,** 1120–1123 (2002).
8. Grime, J. P. *Plant Strategies and Vegetation Processes* (Wiley, Chichester, 1979).
9. Lindeman, R. L. The trophic-dynamic aspect of ecology. *Ecology* **23,** 399–418 (1942).
10. Oksanen, L., Fretwell, S. D., Arruda, J. & Niemelä, P. Exploitation ecosystems in gradients of primary productivity. *Am. Nat.* **118,** 240–261 (1981).
11. Kaunzinger, C. M. K. & Morin, P. J. Productivity controls food-chain properties in microbial communities. *Nature* **395,** 495–497 (1998).
12. Post, D. M., Pace, M. L. & Hairston, N. G. Ecosystem size determines food-chain length in lakes. *Nature* **405,** 1047–1049 (2000).
13. Pimm, S. L. & Lawton, J. H. Number of trophic levels in ecological communities. *Nature* **268,** 329–331 (1977).
14. Moore, J. C., Walter, D. E. & Hunt, H. W. Arthropod regulation of micro- and mesobiota in belowground food webs. *Annu. Rev. Entomol.* **33,** 419–439 (1988).
15. de Ruiter, P. C., van Veen, J. A., Moore, J. C., Brussaard, L. & Hunt, H. W. Calculation of nitrogen mineralisation in soil food webs. *Plant Soil* **157,** 263–273 (1993).
16. Gardner, M. R. & Asby, W. R. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature* **228,** 784 (1970).
17. Hunt, H. W. *et al.* The detrital food web in a shortgrass prairie. *Biol. Fertil. Soils* **3,** 57–68 (1987).
18. Hofbauer, J. & Sigmund, K. in *The Theory of Evolution and Dynamical Systems* 204–206 (Cambridge Univ. Press, Cambridge, 1988).
19. Cohen, J. E., Briand, F. & Newman, C. M. *Community Food Webs: Data and Theory* (Springer, Berlin, 1990).
20. Halaj, J. & Wise, D. T. Terrestrial trophic cascades: how much do they trickle? *Am. Nat.* **157,** 262–281 (2001).
21. Finke, D. L. & Denno, R. F. Predator diversity dampens trophic cascades. *Nature* **429,** 407–410 (2004).
22. Hairston, N. G. Sr, Smith, F. E. & Slobodkin, L. B. Community structure, population control, and competition. *Am. Nat.* **142,** 379–411 (1960).
23. Moore, J. C., de Ruiter, P. C. & Hunt, H. W. Influence of productivity on the stability of real and model ecosystems. *Science* **261,** 906–908 (1993).
24. Holt, R. D. & Polis, G. A. A theoretical framework for intraguild predation. *Am. Nat.* **149,** 745–764 (1997).
25. Scheffer, M. & van Nes, E. H. Self-organized similarity, the evolutionary emergence of groups of similar species. *Proc. Natl Acad. Sci. USA* **103,** 6230–6235 (2006).
26. Bascompte, J., Jordano, P. & Olesen, J. M. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* **312,** 431–433 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Biewenga, P. Bolhuis, B. van der Boom, K. Kampen, M. Veninga and W. Willems for assistance in collecting and analysing the soil samples. We thank S. Burgers, J. Krumins and P. Morin for comments on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.-M.N. (anjute@bas.ac.uk).

©2007 Nature Publishing Group

METHODS

Soil layers and methods of sampling. Soil layers in the four developmental stages were: 1, bare soil, mineral layer (C horizon); 2, sparse grasses, mineral layer (C horizon); 3, dense herbaceous vegetation, humus-rich layer (A horizon) of 2–5 cm; and 4, woody vegetation with undergrowth, organic litter layer (F horizon) of more than 5 cm. Dominating plant species in the above-ground vegetation along the gradient of Schiermonnikoog were: 2, *Festuca rubra*; 3, *Ammophila arenaria* and *F. rubra*; 4, *Sambucus nigra* and *Urtica dioica*. Dominating plant species in the above-ground vegetation along the gradient of Hulshorsterzand were: 2, bryophytes with dispersed *Corynephorus canescens*; 3, mixed vegetation of bryophytes, lichens, *C. canescens* and young *Pinus sylvestris*; 4, *P. sylvestris* forest with an understory of *Deschampsia flexuosa* and bryophytes. The ages of the stages in terms of soil development were approximately 0, 10, 25 and 100 yr in Schiermonnikoog and 0, 5, 15 and 50 yr in Hulshorsterzand.

In each of the four stages in both successions, soil samples were taken three times within a year, in four replications, to establish yearly averages: in November 1993, March 1994 and August 1994 in Schiermonnikoog, and in October 1993, March 1994 and September 1994 in Hulshorsterzand. A total of 96 soil samples were processed as described in refs 15, 27 and 28. Biomass of bacteria was determined using the method from ref. 29, of fungi using ref. 28, and of protozoa, nematodes and microarthropods using ref. 27. Yearly average biomass densities of all functional groups were determined for 32 food webs (Supplementary Information). All measurements for a single food-web replicate were taken from the same soil sample; that is, they represented replications of entire food webs.

Earlier studies on the above-ground vegetation in both series indicate that above-ground productivity increases with age in these stages[5,6]. Observed overall soil nitrogen mineralization rates indicated a similar increase in below-ground productivity along both gradients: 1.9 (1.1–2.9), 7.9 (6.2–9.1), 39 (35–46) and 80 (70–92) kg N ha^{-1} yr^{-1} cm $depth^{-1}$ in Schiermonnikoog, and 1.1 (0.49–2.5), 7.9 (5.3–11.4), 40 (34–48) and 86 (49–160) kg N ha^{-1} yr^{-1} cm $depth^{-1}$ in Hulshorsterzand.

Calculation of interaction strengths. Community matrices were derived using the method from ref. 1. The community matrix A is the jacobian matrix of interaction obtained by linearizing the nonlinear dynamical systems of growth equations and evaluated at the equilibrium with all groups present[1]. The elements of A, partial derivatives of the growth equations, are the interaction strengths α_{ij}, representing the effects of the trophic groups on each other near equilibrium. Assuming direct dependence of feeding rates on predator population density, effects of predator j on prey i are defined as predation per predator biomass in equilibrium: $\alpha_{ij} = -f_{ij}$, where f_{ij} is a mass-specific predation rate of predator j on its prey i in equilibrium[2,4,7,13]. Assuming also direct dependence of feeding rates on prey population density, effects of prey on their predator are defined as predator growth resulting from this predation, per prey biomass in equilibrium. This means that effect sizes of prey i on their predator j were related to effects of the predator on the prey according to: $\alpha_{ji} = -e_{ij}\frac{B_i}{B_j}\alpha_{ij}$, where B_j is the equilibrium biomass of predator j and e_{ij} is the efficiency with which j converts food into biomass, from feeding on i, $0 < e_{ij} < 1$ (refs 2, 4, 7, 13). The same analysis can be performed for descriptions of interaction that incorporate more complex functional responses. This does not change the effects of predators on prey and will not fundamentally change the relationship with effects of prey on predators; it will generally, at most, introduce another stabilizing factor. Thus we were able to concentrate on the simple case outlined above without clouding the issue by more complicated formulations, for which we had little empirical information.

Values of interaction strengths of all the interactions in the soil food webs were calculated using the method of ref. 2. This involved the following procedure. Mass-specific predation rates were derived from observed average biomass densities using the steady state assumption, following the method from ref. 17: $\sum_{i\in\Delta_j} f_{ij}B_j = \frac{d_jB_j + \sum_{k\in\Gamma_j} f_{jk}B_k}{e_j}$, where Δ_j is the set of indices referring to the prey species of j, Γ_j is the set of indices referring to the predator species of j, B_j is the observed yearly average population density of the predator, d_j is the mass-specific natural death rate (per yr), e_j is a conversion efficiency, $0 < e_{ij} < 1$, making the simplifying assumption $e_{ij} = e_j$ for all $i\in\Delta_j$, and using $f_{ij} = \frac{p_{ij}B_i}{\sum_{h\in\Delta_j} p_{hj}B_h} f_j$, where $f_j = \sum_{i\in\Delta_j} f_{ij}$ and p_{ij} is a prey preference factor. Preference factors were published values on the basis of experiments involving feeding trials and gut content analyses in relation to prey availability[17]. The values for these preferences are rough estimates, but they give approximations in terms of orders of magnitude. All parameters are defined as positive.

The diagonal terms in a community matrix referring to self-damping (intraspecific competition) of the organisms were determined on the basis of the empirical values of natural death rates, representing all non-predatory losses that can be expected in populations in their natural environment[17]. Using the method from ref. 2 they were modelled as $\alpha_{ii} = -sd_i$, where d_i is the total natural specific death rate of i in equilibrium (defined positive), and s is the proportion of natural death in equilibrium (that is, assumed equal for all species in a food web) that is attributed to density-dependent mechanisms ($0 < s < 1$; for a detailed argument, see Supplementary Information). The value of s was unspecified and was derived to assess stability (see below).

Calculation of stability. To allow for variation in the matrix element values, 100 community matrices were used to represent each web, drawing off-diagonal element values randomly from uniform intervals around the calculated values $\langle 0, 2\alpha_{ij}\rangle$, $i \neq j$ using the method from ref. 2. Departing from ref. 2 and using the method from ref. 7, diagonal elements $\alpha_{ii} = -sd_i$ were not set at certain fixed, arbitrary levels, but were varied. The procedure is as follows: matrix stability was determined in the standard way, that is, by calculating the eigenvalues; if all eigenvalues have negative real parts, a matrix is stable. Values for s were assessed as the minimum value that was necessary for matrix stability (for details of the procedure, see Supplementary Information). Stability of a food web was then calculated as the mean value of s of the 100 sampled community matrices. This provided a robust measure of stability. For further explanation and motivation, see Supplementary Information.

Owing to a lack of detailed biological information, we have chosen the self-damping factor s to be the same for all species and to be constant during successional development. We acknowledge that the levels of intra-specific interaction could vary between the species in a food web and change over succession for independent reasons. Also, we note that stability was not measured directly by, for example, perturbing food webs experimentally. It remains to be determined how good a surrogate our measure is. However, it should be noted that such experiments would be difficult because 'stability' is not a directly measurable variable and, even with experiments, surrogate quantities would have to be used.

Calculation of average interaction strength. Average interaction strengths were calculated:

$$a = \frac{\left\{\left|\sum_{j\in M}\sum_{i\in\Delta_j}\alpha_{ij}\right|\left(\sum_{j\in M}\sum_{i\in\Delta_j}\alpha_{ji}\right)\right\}^{\frac{1}{2}}}{L(d_1\ldots d_n)^{\frac{1}{n}}}$$

where L is the total number of feeding links, M is the set of indices referring to all predator species, Δ_j is the set of indices referring to the prey species of j, and d_i—the total natural specific death rate of i in equilibrium (defined positive)—is a scaling factor used to make this metric comparable with the stability metric, which is also relative to these death rates. The geometric means of negative top-down and positive bottom-up effects were taken separately because negative and positive interaction strengths differed systematically by about two orders of magnitude[2]. The precise way in which a mean was taken had no effect on the resulting patterns.

Calculation of loop weights. Loop weight is defined as the geometric mean of the absolute values of the jacobian elements in a loop[7]:

$$w^{(k)} = \left|\frac{\alpha_{i_1i_2}\alpha_{i_2i_3}\ldots\alpha_{i_ki_1}}{d_{i_1}d_{i_2}\cdots d_{i_k}}\right|^{1/k}$$

where $w^{(k)}$ is the scaled weight of a loop of length k. There can be different loops of length k, with consequently different loop weights.

Parameterization of 'Cohen webs'. Community matrices were derived for 95 food webs assembled in ref. 19 ('Cohen webs') from hypothetical biomass structures of a 10-fold and 20-fold biomass decrease per trophic level, using the method from ref. 7. They were parameterized with equal values for all groups, assuming $e_{ij} = e = 0.1$, $d_j = d = 1$, and no specific prey preferences.

27. Brussaard, L., Bouwman, L. A., Geurs, M., Hassink, J. & Zwart, K. B. Biomass, composition and temporal dynamics of soil organisms of a silt loam soil under conventional and integrated management. *Neth. J. Agr. Sci.* **38,** 283–302 (1990).
28. Hassink, J., Lebbink, G. & van Veen, J. A. Microbial biomass and activity of a reclaimed polder soil under a conventional or a reduced-input farming system. *Soil Biol. Biochem.* **23,** 515–524 (1991).
29. Bloem, J., Veninga, M. & Shepherd, J. Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. *Appl. Environ. Microb.* **61,** 926–936 (1995).

©2007 Nature Publishing Group

ADDENDUM

doi:10.1038/nature07825

Reconciling complexity with stability in naturally assembling food webs

Anje-Margriet Neutel, Johan A. P. Heesterbeek, Johan van de Koppel, Guido Hoenderboom, An Vos, Coen Kaldeway, Frank Berendse & Peter C. de Ruiter

Nature 449, 599–602 (2007)

In the Supplementary Information of this Letter, no biomass data were given for the two basal food-web compartments: roots and detritus. These were not measured in the soil samples (see also ref. 2). We used values for root biomass of 900 kg C ha^{-1} cm $depth^{-1}$ in all webs with roots present. We used values for detritus biomass of 4, 25, 250 and 2,500 kg C ha^{-1} cm $depth^{-1}$ for the successional stages 1 to 4, respectively, of both series (roughly corresponding to the increase in organic matter along the productivity gradients). Other than in the one replicate in stage 1 (Schiermonnikoog) where phytophageous nematodes were present (for which basal values of stage 2 were used), we assumed the same values for the replicates within each stage. The choice of these particular values does not affect our key findings. The data to do the stability analyses are provided in the Letter, its Supplementary Information or via refs 2 and 17. The estimated physiological parameter values, the matrix of feeding relations and prey preferences, formalization of detritus feedbacks and an example (Jacobian) community matrix are provided in the Supplementary Information to this Addendum.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Light adaptation in cone vision involves switching between receptor and post-receptor sites

Felice A. Dunn[1], Martin J. Lankheet[4] & Fred Rieke[2,3]

We see over an enormous range of mean light levels, greater than the range of output signals retinal neurons can produce. Even highlights and shadows within a single visual scene can differ ~10,000-fold in intensity—exceeding the range of distinct neural signals by a factor of ~100. The effectiveness of daylight vision under these conditions relies on at least two retinal mechanisms that adjust sensitivity in the ~200 ms intervals between saccades[1]. One mechanism is in the cone photoreceptors (receptor adaptation)[2–5] and the other is at a previously unknown location within the retinal circuitry that benefits from convergence of signals from multiple cones (post-receptor adaptation)[6,7]. Here we find that post-receptor adaptation occurs as signals are relayed from cone bipolar cells to ganglion cells. Furthermore, we find that the two adaptive mechanisms are essentially mutually exclusive: as light levels increase the main site of adaptation switches from the circuitry to the cones. These findings help explain how human cone vision encodes everyday scenes, and, more generally, how sensory systems handle the challenges posed by a diverse physical environment.

Adaptation enables weak sensory inputs to be amplified and protected from noise while preventing strong inputs from saturating neural responses. Past physiological work has shown that ganglion cells[6,7], the output cells of the retina, adapt at lower light levels than cones[2–5] and one of the downstream targets of the cones, horizontal cells[8–10]. Such studies provide evidence for adaptation in the retinal circuitry and in the cone photoreceptors. However, differences in recording conditions hinder direct comparison of cone and ganglion cell adaptation; thus, the properties of receptor and post-receptor adaptation remain unclear. In addition, past studies have not separated cone- and rod-mediated signals, and rods or rod circuitry could contribute to the identified post-receptor adaptation (see Supplementary Information). Here we track cone-mediated signals from the input to the output cells of the primate retina to identify where adaptation occurs and to compare directly the properties of adaptation at different sites.

We studied the midget (parvocellular-projecting) and parasol (magnocellular-projecting) circuits that form two key parallel readouts of the cone signals (Fig. 1a). Figure 1 characterizes adaptation of cone photoreceptors and cone-mediated currents of ON (depolarizing to light increments) ganglion cells measured in directly comparable conditions (cones in slices, ganglion cells in a flat mount; see Methods). The Methods and Supplementary Information describe isolation of cone-mediated signals (Supplementary Figs 1 and 2).

Both receptor and post-receptor adaptation were apparent. Figure 1c, d shows responses to brief flashes in the absence and presence of adapting backgrounds with intensities near typical interior lighting (Fig. 1c;

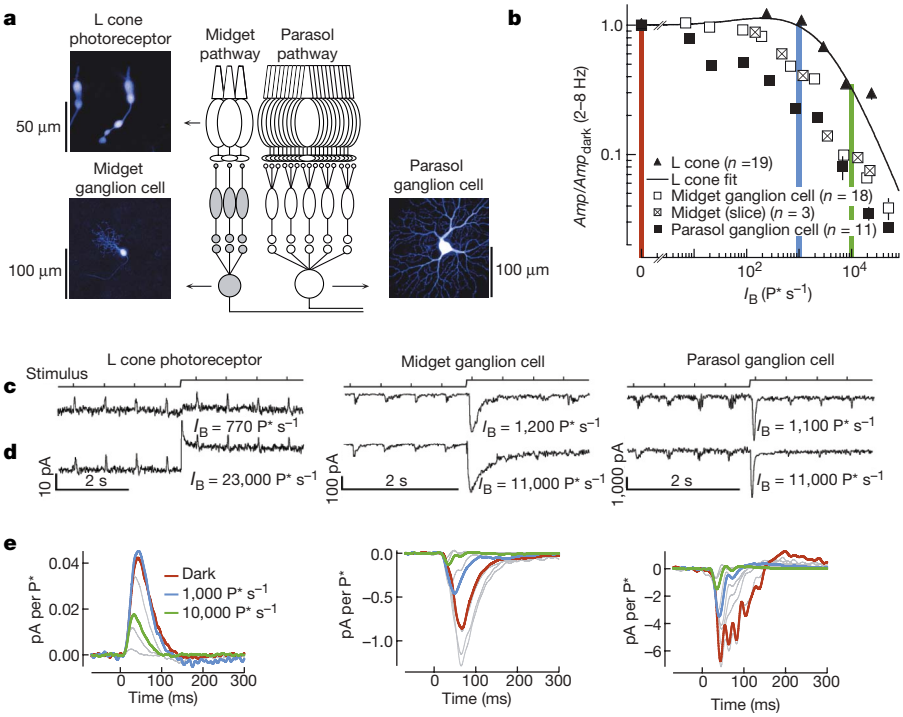


Figure 1 | Midget and parasol ganglion cells adapt at lower backgrounds than L cones. **a**, Schematic of primate midget and parasol pathways with fluorescent images of cones in slice and ganglion cells in flat mount. **b**, Background dependence of the response amplitude (mean ± s.e.m.). Amplitudes (*Amp*) were measured between 2–8 Hz and normalized by the amplitude in darkness (Amp_{dark}). Cone data were fitted from equation (1) (Methods) with $a = 100$, $b = 1.3$ and $c = 0.00029$. **c**, **d**, Current responses to flashes in darkness and on backgrounds producing ~1,000 and ~10,000 P* s^{-1}. **e**, Average cone current responses (in pA) normalized by the flash strength (in P*) in darkness (red) and on backgrounds producing 1,100 (blue) and 7,600 (green) P* s^{-1}, responses of ON midget ganglion cells on backgrounds producing 0.5 (red), 690 (blue) and 7,000 (green) P* s^{-1}, and responses of ON parasol ganglion cells on backgrounds producing 0.5 (red), 870 (blue) and 6,800 (green) P* s^{-1}. Grey traces show responses at intermediate backgrounds.

[1]Program in Neurobiology and Behavior, [2]Howard Hughes Medical Institute and [3]Department of Physiology and Biophysics, University of Washington, Seattle, Washington 98195, USA. [4]Functional Neurobiology and Helmholtz Institute, Utrecht University, 3584 CH Utrecht, The Netherlands.

©2007 Nature Publishing Group

1,000 absorbed photons per cone per second; P* s^{-1}) and indirect sunlight (Fig. 1d; 10,000 P* s^{-1}). Figure 1e collects average flash responses from multiple cells of each type at each background tested. Dim backgrounds (less than 1,000 P* s^{-1}) produced a small increase in the cone flash responses (compare red and blue traces), but substantially reduced the amplitude and duration of the responses of both midget and parasol ganglion cells. Thus post-receptor adaptation dominated at low backgrounds. Bright backgrounds (greater than 10,000 P* s^{-1}) reduced the amplitude and duration of the responses of cones and ganglion cells (compare blue and green traces). The cone responses continued to adapt without circuit feedback (see Supplementary Fig. 4). Previous primate cone recordings did not exhibit substantial changes in kinetics with background light[4,5]; the background-dependent kinetic changes we find suggest that primate cones are similar to cones of other species[11,12], rather than the anomaly (see Supplementary Information).

Ganglion cell responses were reduced by backgrounds 10–100 times dimmer than those that reduced cone responses, consistent with past work[6,7]. Figure 1b summarizes the background dependence of response amplitude in cones and ganglion cells. Amplitudes were averaged for response components between 2 and 8 Hz, because adaptation is strongest at these temporal frequencies[6,7]. Midget ganglion cells recorded in flat mount and slice preparations exhibited similar adaptation, indicating that slicing did not substantially compromise the underlying mechanisms. Parasol ganglion cell responses were reduced more and at lower backgrounds than midget ganglion cell responses, demonstrating that post-receptor adaptation at low light levels was produced independently in the two pathways.

To identify the location of post-receptor adaptation, we recorded from the ON midget and DB4 or DB5 cone bipolar cells that convey cone signals to midget and parasol ganglion cells (Fig. 2a)[13]. To test for adaptation of the synapse between cones and cone bipolar cells, we held the cone bipolar voltage fixed and measured the light-evoked input currents. As with the cones, modest backgrounds (Fig. 2b, blue traces in Fig. 2d) produced little or no adaptation in the cone bipolar currents

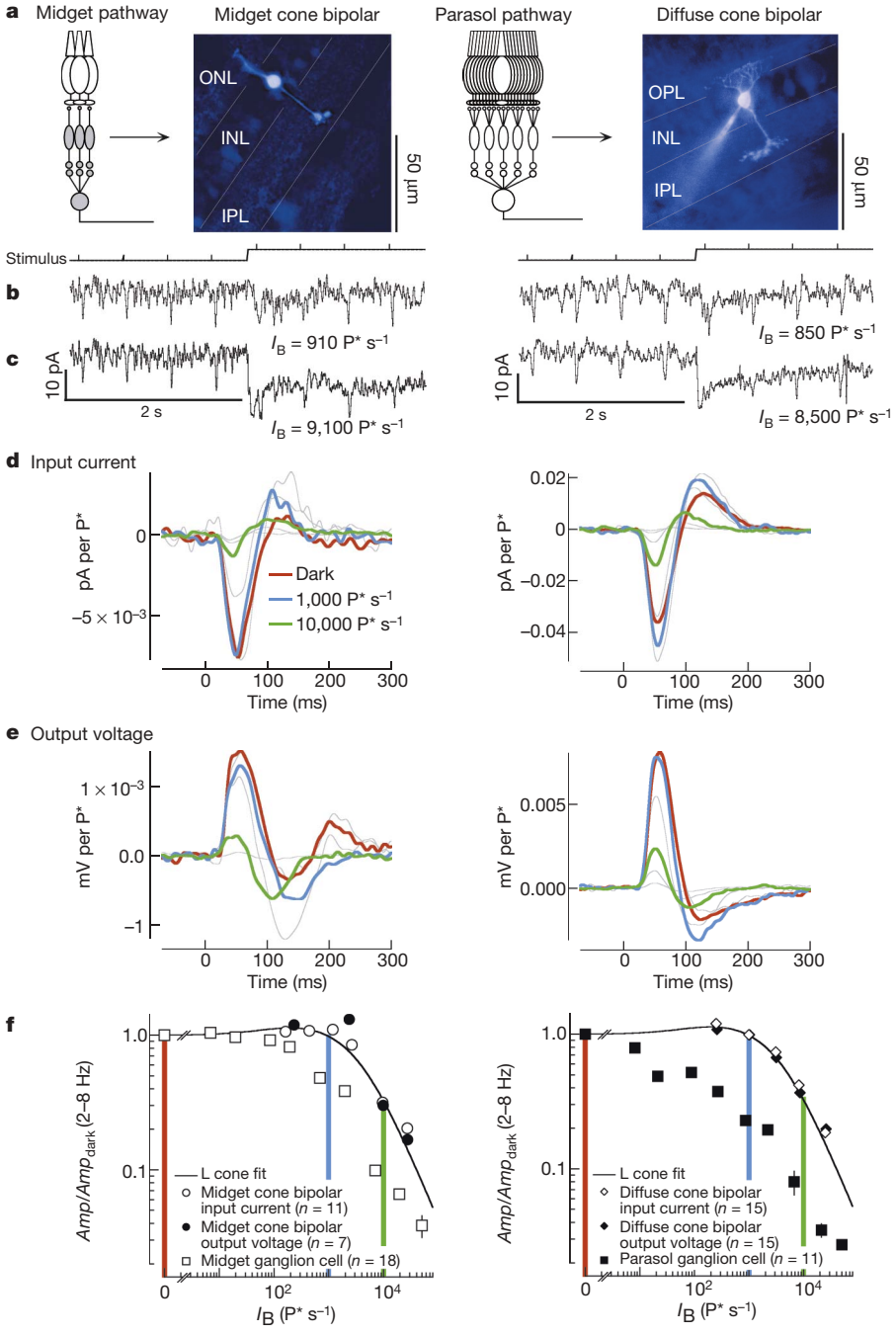


Figure 2 | Post-receptor adaptation occurs in signal transfer from cone bipolar cells to ganglion cells. **a**, Fluorescent images of a midget cone bipolar cell (left panel) and a diffuse cone bipolar cell (right panel) in slice. OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer. **b**, **c**, Current responses to flashes in darkness and on backgrounds producing ~1,000 and ~10,000 P* s^{-1}. **d**, **e**, Left panels, average current (**d**) and voltage (**e**) responses of midget cone bipolar cells in darkness (red) and on backgrounds producing 1,200 (870 for voltage analysis) P* s^{-1} (blue) and 9,600 (9,900) P* s^{-1} (green). Right panels, average current (**d**) or voltage (**e**) responses of diffuse cone bipolar cells in darkness (red) and on backgrounds producing 1,000 (1,000 for voltage analysis) P*s^{-1} (blue) and 8,200 (8,600) P* s^{-1} (green). Grey traces show responses at intermediate backgrounds. **f**, Background dependence of response amplitude (mean ± s.e.m.). Amplitudes (*Amp*) were measured between 2–8 Hz and normalized by the amplitude in darkness (Amp_{dark}). Cone fit and ganglion cells are from Fig. 1b.

©2007 Nature Publishing Group

whereas bright backgrounds (Fig. 2c, green traces in Fig. 2d) produced substantial adaptation. Figure 2f compares the background dependence of the amplitude of the cone bipolar current responses with that of the cones and ganglion cells; adaptation in both the midget and diffuse cone bipolar currents tracked that of the cones and showed no evidence for the post-receptor adaptation apparent in the ganglion cell responses.

To test for adaptation in the conversion of cone bipolar synaptic currents to voltages, we characterized the effect of backgrounds on the bipolar voltage responses. The background dependence of the cone bipolar voltages (Fig. 2e, f) resembled that of their synaptic input currents and of the cone signals. The absence of post-receptor adaptation in the output voltages of midget and diffuse cone bipolar cells, and its presence in the midget and parasol ganglion cell synaptic inputs, localizes post-receptor adaptation to signal transfer between cone bipolar cell and ganglion cell in both pathways. This location has been implicated in other forms of adaptation[14–16].

How do receptor and post-receptor adaptation together control sensitivity? The relative contributions of receptor and post-receptor adaptation were clearer when individual temporal frequencies of the responses were examined in the frequency domain. The left column of Fig. 3 compares amplitude spectra for cones, diffuse cone bipolar cells and parasol ganglion cells in darkness and in the presence of a background producing 1,000 P* s^{-1}—spanning the range of backgrounds where post-receptor adaptation dominates and the cones adapt little,

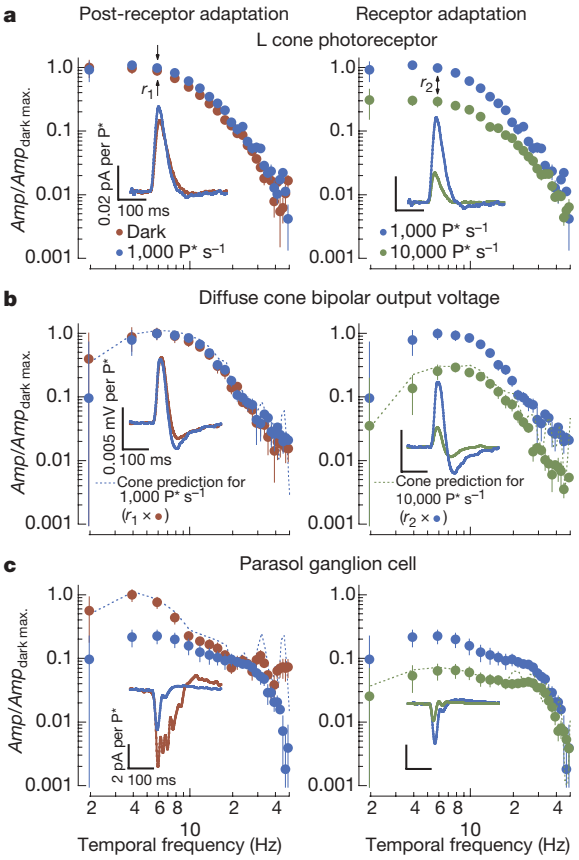


Figure 3 | The primary site of adaptation switches from circuitry to cones as light levels increase. **a**, **b**, **c**, Amplitude spectra of flash responses for L cones (**a**), diffuse cone bipolar cells (**b**) and parasol ganglion cells (**c**). Amplitudes (*Amp*) at each temporal frequency were normalized by the maximum amplitude spectra in the dark ($Amp_{dark\ max.}$). Left panels, responses in darkness (red) and in the presence of background producing ~1,000 P* s^{-1} (blue); right panels, responses in the presence of backgrounds producing ~1,000 (blue) and ~10,000 P* s^{-1} (green). Insets show time-domain responses. Adaptation predicted from the cone responses (dotted lines in **b** and **c**) was calculated by multiplying the postsynaptic amplitude spectra at the lower of the two backgrounds by the appropriate ratio of the cone amplitude spectra (r_1, r_2). Data are mean ± s.e.m.

if at all. The right panels compare amplitude spectra for backgrounds producing 1,000 and 10,000 P* s^{-1}—a range throughout which cones contribute to adaptation. The insets show time-domain flash responses. The amplitude spectra allowed us to determine how much adaptation in the cone bipolar cells and ganglion cells was inherited from the cones themselves and how much was produced by post-receptor adaptation. We applied the ratio of cone amplitude spectra at two backgrounds (for example, ratio of the red to blue points in Fig. 3a, left panel) to the amplitude spectrum of the downstream cell at the lower background (for example, blue points in Fig. 3b, left panel). This procedure generated a predicted amplitude spectrum (for example, blue dotted line in Fig. 3b, left panel) for the downstream cell, assuming adaptation was entirely inherited from that in the cones.

At low backgrounds, the diffuse cone bipolar voltages matched the cone prediction well, consistent with little or no effect of post-receptor adaptation, whereas the parasol ganglion cell responses differed substantially from the cone prediction because of post-receptor adaptation (Fig. 3c, left panel). The contributions of cone and downstream adaptation were quite different when the cones began to adapt (Fig. 3, right panel). The change in diffuse cone bipolar cell and parasol ganglion cell spectra between 1,000 and 10,000 P* s^{-1} matched the predictions based on cone adaptation (dashed green lines in Fig. 3b, c, right panel), indicating that cone adaptation dominated adaptation of the bipolar and ganglion cell responses produced by the change in background. Thus a single site dominated adaptation at all backgrounds: post-receptor adaptation at low backgrounds and receptor adaptation at high backgrounds.

How do the properties of receptor and post-receptor adaptation compare? Figure 3 shows clear differences in how the two forms of adaptation affected ganglion cell responses. Post-receptor adaptation preferentially attenuated low frequencies (Fig. 3c, left), consistent with previous studies[6,7,17], whereas receptor adaptation attenuated all temporal frequencies (Fig. 3c, right; Supplementary Fig. 3; see also ref. 5). Thus the dominant form of adaptation depends on both light intensity and temporal frequency; these dependencies are likely to contribute to inconsistencies in previous descriptions of adaptation based on behaviour (reviewed by ref. 18).

Having distinct operating ranges requires coordination to cease post-receptor adaptation when receptor adaptation becomes active. The dominance of cone adaptation at high backgrounds cannot be explained simply by cone adaptation eliminating the signal producing post-receptor adaptation, because cones continue to produce distinguishable maintained currents at backgrounds well above those where post-receptor adaptation saturates (Fig. 4a; see also figure 7 in ref. 4). Instead, at high backgrounds post-receptor adaptation must become insensitive to changes in the mean cone signal, explaining the switch from post-receptor to receptor adaptation.

What rules determine the backgrounds at which receptor and post-receptor adaptation operate? Cone responses to backgrounds where post-receptor adaptation operates are obscured by noise (Fig. 4b, left; ref. 5), particularly when viewed for the ~100 ms period characterizing the integration time for adaptation[10,19] (see also ref. 20). This high level of noise would cause the gain of a mechanism based on the response of a single cone to fluctuate, making adaptation ineffective or even detrimental. Indeed, cone adaptation begins only for backgrounds that produce noticeable changes in the cone response (Fig. 4b, right panel). Thus the onset of adaptation seems to reflect a balance of the benefits of matching a cell's limited response range to the range of input signals encountered[21] and the drawbacks associated with basing such gain controls on noisy signals.

Convergence in the retinal circuitry can improve signal-to-noise, providing the adaptive mechanism with a more reliable signal. Diffuse cone bipolar cells receive input from ~10 cones, but the noise in their flash responses at low backgrounds is similar to that of the cones (Fig. 4c; ratio of cone bipolar to cone signal-to-noise 0.67 ± 0.06; mean ± s.e.m.; see Methods). This is consistent with the lack of post-receptor adaptation in the cone bipolar signals.

©2007 Nature Publishing Group

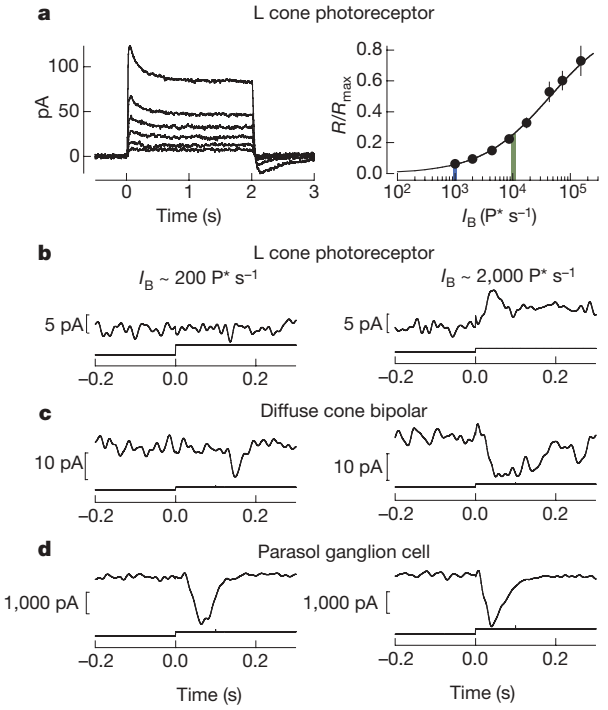


Figure 4 | Signal-to-noise of the adapting signal improves with convergence and increasing light level. a, Left panel, current responses of an L cone to steps doubling in intensity from 4,800 P* s^{-1}; right panel, normalized steady state response amplitudes for 12 L cones fitted with equation (2) (Methods; $I_{1/2} = 45{,}000 \pm 7{,}000$ P* s^{-1}, $n = 0.7 \pm 0.05$). Current responses of an L cone (**b**), a diffuse cone bipolar cell (**c**) and a parasol ganglion cell (**d**) to steps producing 200 P* s^{-1} (left panels) and 2,000 P* s^{-1} (right panels).

The convergence necessary for post-receptor adaptation could be provided by an amacrine cell or by voltage-independent synaptic integration within the ganglion cell. For example, peripheral parasol ganglion cells receive input from hundreds of cones, and their flash responses achieve a higher signal-to-noise ratio at low backgrounds (Fig. 4d; ratio of parasol to cone signal-to-noise, 6.0 ± 0.3). Thus an adaptive mechanism based on converging cone inputs can operate more reliably at low backgrounds than one based on cone or cone bipolar signals. This reasoning suggests that post-receptor adaptation in central midget ganglion cells, which receive input from single cones, could differ considerably from that in the periphery.

As we make saccades to explore a visual scene, retinal neurons encounter a wide range of light intensities. Receptor and post-receptor adaptation permit the amplification required to see objects in shadows while avoiding saturation from the sky. The combination of these adaptive mechanisms allows the visual system to encode details in a scene with greater fidelity than a standard camera at a single exposure setting. The strategy the retina employs—shifting the dominant site of adaptation to match the reliability of the input signals—demonstrates an elegant principle for accurate information processing in sensory perception.

METHODS SUMMARY

Primate (*Macaca fascicularis*, *Macaca nemestrina* and *Papio anubis*) retinas were obtained through the Tissue Distribution Program of the National Primate Research Center at the University of Washington and dark-adapted for 1 h before isolation from the pigment epithelium. Isolated flat-mounted pieces of retina or retinal slices were perfused with bicarbonate-based Ames medium (Sigma) equilibrated with 5% CO_2/95% O_2 at 34–37 °C while we recorded the current or voltage responses of individual cone photoreceptors, cone bipolar cells and ganglion cells. Cells were targeted for recording under infrared differential interference contrast microscopy. Cells were identified at the end of each recording from fluorescence images.

Spatially uniform light stimuli were delivered from three light-emitting diodes (LEDs) with peak outputs at 470, 513 and 640 nm. An accurate view of adaptation of cone-mediated signals required separation of rod- and cone-mediated responses at the lowest backgrounds possible. Cone-mediated signals were probed with brief 640 nm flashes that preferentially stimulated long-wavelength-sensitive cone photoreceptors in the absence and presence of constant background light produced by simultaneous activation of all three LEDs. Contributions from the rod photoreceptors were minimized by pharmacologically blocking rod-mediated inputs to cone bipolar cells or by suppressing rod-mediated signals with short-wavelength-adapting light in cone bipolar and ganglion cells. Responses to brief flashes on varying background lights were averaged across cells and analysed in the time and frequency domains.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 June; accepted 6 August 2007.
Published online 12 September 2007.

1. Yarbus, A. L. *Eye Movements and Vision* (Plenum, New York, 1967).
2. Boynton, R. M. & Whitten, D. N. Visual adaptation in monkey cones: recordings of late receptor potentials. *Science* **170,** 1423–1426 (1970).
3. Valeton, J. M. & van Norren, D. Light adaptation of primate cones: an analysis based on extracellular data. *Vision Res.* **23,** 1539–1547 (1983).
4. Schnapf, J. L., Nunn, B. J., Meister, M. & Baylor, D. A. Visual transduction in cones of the monkey Macaca fascicularis. *J. Physiol. (Lond.)* **427,** 681–713 (1990).
5. Schneeweis, D. M. & Schnapf, J. L. The photovoltage of macaque cone photoreceptors: adaptation, noise, and kinetics. *J. Neurosci.* **19,** 1203–1216 (1999).
6. Lee, B. B., Pokorny, J., Smith, V. C., Martin, P. R. & Valberg, A. Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *J. Opt. Soc. Am. A* **7,** 2223–2236 (1990).
7. Purpura, K., Tranchina, D., Kaplan, E. & Shapley, R. M. Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis. Neurosci.* **4,** 75–93 (1990).
8. Lee, B. B., Dacey, D. M., Smith, V. C. & Pokorny, J. Horizontal cells reveal cone type-specific adaptation in primate retina. *Proc. Natl Acad. Sci. USA* **96,** 14611–14616 (1999).
9. Smith, V. C., Pokorny, J., Lee, B. B. & Dacey, D. M. Primate horizontal cell dynamics: an analysis of sensitivity regulation in the outer retina. *J. Neurophysiol.* **85,** 545–558 (2001).
10. Lee, B. B., Dacey, D. M., Smith, V. C. & Pokorny, J. Dynamics of sensitivity regulation in primate outer retina: the horizontal cell network. *J. Vis.* **3,** 513–526 (2003).
11. Baylor, D. A. & Hodgkin, A. L. Changes in time scale and sensitivity in turtle photoreceptors. *J. Physiol. (Lond.)* **242,** 729–758 (1974).
12. Matthews, H. R., Fain, G. L., Murphy, R. L. & Lamb, T. D. Light adaptation in cone photoreceptors of the salamander: a role for cytoplasmic calcium. *J. Physiol. (Lond.)* **420,** 447–469 (1990).
13. Boycott, B. B. & Wassle, H. Morphological classification of bipolar cells of the primate retina. *Eur. J. Neurosci.* **3,** 1069–1088 (1991).
14. Baccus, S. A. & Meister, M. Fast and slow contrast adaptation in retinal circuitry. *Neuron* **36,** 909–919 (2002).
15. Hosoya, T., Baccus, S. A. & Meister, M. Dynamic predictive coding by the retina. *Nature* **436,** 71–77 (2005).
16. Manookin, M. B. & Demb, J. B. Presynaptic mechanism for slow contrast adaptation in mammalian retinal ganglion cells. *Neuron* **50,** 453–464 (2006).
17. Tranchina, D., Gordon, J. & Shapley, R. M. Retinal light adaptation—evidence for a feedback mechanism. *Nature* **310,** 314–316 (1984).
18. Hood, D. C. Lower-level visual processing and models of light adaptation. *Annu. Rev. Psychol.* **49,** 503–535 (1998).
19. Yeh, T., Lee, B. B. & Kremers, J. The time course of adaptation in macaque retinal ganglion cells. *Vision Res.* **36,** 913–931 (1996).
20. Fairhall, A. L., Lewen, G. D., Bialek, W. & de Ruyter Van Steveninck, R. R. Efficiency and ambiguity in an adaptive neural code. *Nature* **412,** 787–792 (2001).
21. Laughlin, S. A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. C* **36,** 910–912 (1981).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Dacey, O. Packer, B. Peterson and T. Haun for providing primate tissue; P. Newman and M. Wixey for technical assistance; J. Cafaro, E. J. Chichilnisky, T. Doan, W. Dunn, G. Horwitz, R. Kiani, G. Murphy, F. Soo, B. Wark for comments on the manuscript; and B. Lundstrom for comments on figures. This work was supported by the National Institutes of Health (F.R.), Achievement Rewards for College Scientists Foundation (F.A.D.) and the Howard Hughes Medical Institute (F.R., and F.A.D. with a predoctoral fellowship).

Author Contributions F.A.D., M.J.L. and F.R. participated in all aspects of this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.R. (rieke@u.washington.edu).

©2007 Nature Publishing Group

METHODS

Tissue. Slice and flat mount preparations of primate retina followed procedures described previously[22]. All recordings were from peripheral retina (>30° eccentricity, based on ganglion cell dendritic field sizes) isolated from the pigment epithelium.

Recordings and cell types. Cone photoreceptors were recorded in slices prepared as described previously[23,24]. Currents (holding potential −60 mV) were measured using perforated or whole-cell voltage clamp with an internal solution containing (in mM): 125 potassium-aspartate, 10 KCl, 10 HEPES, 5 NMG-HEDTA, 1 $MgCl_2$, 0.5 $CaCl_2$, 4 Mg-ATP, 0.5 Tris-GTP, 0.1 Alexa-488 or Alexa-750 and 0.05 mg ml^{-1} amphotericin-B for perforated-patch recordings; pH was adjusted to 7.2 with NMG-OH and osmolarity was ~280 mOsm. Long (L)-wavelength-sensitive cones were distinguished from middle (M)- and short (S)-wavelength-sensitive cones on the basis of their higher sensitivity to 640 nm light.

Cone bipolar responses were measured in slices by whole-cell voltage- or current-clamp recordings with the same internal solution used for the cones. Each recorded cell was identified as a diffuse or midget cone bipolar from fluorescence images acquired after the experiment (Fig. 2a)[13]. Ganglion cells were recorded in a flat mount preparation using an internal solution containing (in mM): 105 $CsCH_3SO_3$, 10 TEA-Cl, 20 HEPES, 10 Cs_2-EGTA, 2 QX-314, 5 Mg-ATP, 0.5 Tris-GTP and 0.1 Alexa-488 or Alexa-750; pH was adjusted to 7.2 with Cs-OH and osmolarity was ~280 mOsm. Midget and parasol cells were targeted on the basis of soma size and sustained versus transient responses to light steps. Identification was confirmed with fluorescence images acquired after each recording (Fig. 1a). We restricted our study to ON cone bipolar and ON ganglion cells. Junction potentials (approximately −10 mV) have not been corrected.

Light stimuli and isolation of cone-mediated responses. All stimuli uniformly illuminated a 630 μm diameter circular area centred on the recorded cell. Calibrated photon fluxes (photons μm^{-2} s^{-1}) were converted to absorbed photons per cone per second (P* s^{-1}) on the basis of the L cone spectral sensitivity, measured LED spectrum, and a collecting area of 0.37 μm^2 (ref. 4).

Adapting backgrounds generated by simultaneous activation of all three LEDs produced near equal photon absorption rates in L and M cones, and eightfold lower rates in S cones. Responses with and without backgrounds were probed with 10-ms flashes from the 640 nm LED. Flash strengths were chosen to elicit responses with amplitudes < 25% of the maximal response. In a subset of experiments we used multiple flash strengths to check for linearity; none of the results reported depended systematically on flash strength. Flashes from the 640 nm LED stimulated L cones 5 times more effectively than M cones, whereas the adapting background equally stimulated L and M cones. The combination of flash and background minimized the possibility that M cones could compensate for the L cones in downstream retinal cells that receive inputs from both cone types. M and L cones adapted similarly (data not shown).

We probed cellular responses using long-wavelength light to isolate cone inputs with minimal rod contamination (Supplementary Fig. 1). Light from the 640 nm LED was absorbed 560 times more effectively by L cones than rods. In the cones themselves this preferential activation was sufficient to eliminate measurable contamination from electrically-coupled rods (Supplementary Fig. 1a)[25]. However, cone bipolar cell and ganglion cell responses to 640-nm wavelength flashes delivered in the dark still had a substantial rod-mediated component. For ganglion cells, rod signals were suppressed by a constant background from the 470 nm LED producing ~3 (for parasol cells) and up to 120 (for midget cells) absorbed photons per rod per second (Supplementary Fig. 1b, c, bottom); this background reduced the rod responses by a factor of at least 10 without noticeably affecting cone responses (Supplementary Fig. 1b, c, top). For ON cone bipolar cells, signals mediated by the rod bipolar pathway were suppressed by blocking rod bipolar cell input to AII amacrine cells with the AMPA-receptor antagonist DNQX (10 μM) or NBQX (5 μM). Similar results were obtained when rod signals were suppressed by 470 nm light (Supplementary Fig. 2).

Data processing. Data were filtered at a bandwidth of 0–60 Hz in Figs 1c, d, and 2b, c, and at a bandwidth of 0–100 Hz in Figs 1e and 2d, e. The background dependence of the response amplitude was taken as the average of the amplitude spectra across 2–8 Hz in Figs 1b and 2f, and across 18–22 Hz in Supplementary Fig. 3.

Fits. The dependence of response amplitude on background (Figs 1b and 2f) were fitted empirically:

$$fit = \left(\frac{|a| + bI_B}{|a| + I_B}\right)(cI_B + 1)^{-1} \quad (1)$$

The first term accounts for the increase in response amplitude with background seen in some cells: a, determines the intensity at which the increase begins and b the amount of increase. The second term describes the decreased amplitude with background, with c setting when the decrease begins.

In Fig. 4a, the background dependence of steady state response amplitude normalized by the maximum current, R/R_{max}, was fitted with the Hill equation:

$$\frac{R}{R_{max}} = \frac{1}{1 + (I_{1/2}/I_B)^n} \quad (2)$$

where $I_{1/2}$ is the half maximal background and n is the Hill exponent describing the relationship between response amplitude and background intensity, I_B.

Signal-to-noise. Signal-to-noise was calculated as the ratio between the flash response amplitude scaled by the number of isomerizations (for example, red traces in Fig. 1e and 2d) and the standard deviation of current in darkness.

Image processing. Output levels, brightness and contrast of background and cell images in Figs 1a and 2a were adjusted independently to enhance the cell profile (Adobe Photoshop). Background image of retinal layers in Fig. 2a was obtained from extracellular sulphurhodamine. Each panel is a projection of a stack of confocal images.

22. Dunn, F. A., Doan, T., Sampath, A. P. & Rieke, F. Controlling the gain of rod-mediated signals in the mammalian retina. *J. Neurosci.* **26,** 3959–3970 (2006).
23. Field, G. D. & Rieke, F. Nonlinear signal transfer from mouse rods to bipolar cells and implications for visual sensitivity. *Neuron* **34,** 773–785 (2002).
24. Armstrong-Gold, C. E. & Rieke, F. Bandpass filtering at the rod to second-order cell synapse in salamander (Ambystoma tigrinum) retina. *J. Neurosci.* **23,** 3796–3806 (2003).
25. Schneeweis, D. M. & Schnapf, J. L. Photovoltage of rods and cones in the macaque retina. *Science* **268,** 1053–1056 (1995).

©2007 Nature Publishing Group

LETTERS

Inhibition of nociceptors by TRPV1-mediated entry of impermeant sodium channel blockers

Alexander M. Binshtok[1], Bruce P. Bean[2] & Clifford J. Woolf[1]

Most local anaesthetics used clinically are relatively hydrophobic molecules that gain access to their blocking site on the sodium channel by diffusing into or through the cell membrane[1]. These anaesthetics block sodium channels and thereby the excitability of all neurons, not just sensory neurons. We tested the possibility of selectively blocking the excitability of primary sensory nociceptor (pain-sensing) neurons by introducing the charged, membrane-impermeant lidocaine derivative QX-314 through the pore of the noxious-heat-sensitive TRPV1 channel. Here we show that charged sodium-channel blockers can be targeted into nociceptors by the application of TRPV1 agonists to produce a pain-specific local anaesthesia. QX-314 applied externally had no effect on the activity of sodium channels in small sensory neurons when applied alone, but when applied in the presence of the TRPV1 agonist capsaicin, QX-314 blocked sodium channels and inhibited excitability. Inhibition by co-applied QX-314 and capsaicin was restricted to neurons expressing TRPV1. Injection of QX-314 together with capsaicin into rat hindpaws produced a long-lasting (more than 2 h) increase in mechanical and thermal nociceptive thresholds. Long-lasting decreases in pain sensitivity were also seen with regional injection of QX-314 and capsaicin near the sciatic nerve; however, in contrast to the effect of lidocaine, the application of QX-314 and capsaicin together was not accompanied by motor or tactile deficits.

Although the goal of topical or regional anaesthesia is to block the transmission of signals in nociceptors to prevent pain, the administration of local anaesthetics also produces numbness from block of low-threshold pressure and touch receptors, paralysis from block of motor axons, and block of autonomic fibres. Nociceptor neurons are unique in expressing high-threshold transducer channels, including, in most, the TRPV1 receptor, which forms ion channels activated by both noxious heat and capsaicin, the pungent ingredient in chilli peppers[2]. The pore of TRPV1 channels is suggested to be large enough to pass compounds at least as large as FM1-43, a dye of molecular mass 452 Da (ref. 3). QX-314, a positively charged derivative of lidocaine with a molecular mass of 263 Da, has no effect on neuronal sodium channels when applied extracellularly but does block sodium channels when applied intracellularly[4–8]. This indicated to us that it might be possible to use the TRPV1 channel to deliver QX-314 or other charged channel blockers selectively into nociceptor neurons by applying such blockers together with TRPV1 agonists.

We recorded current through voltage-dependent sodium channels by using whole-cell voltage-clamp recordings from adult rat dorsal root ganglion (DRG) neurons and identified nociceptors by responsiveness to a short (1-s) application of 1 μM capsaicin. In all small (24 ± 5 μm (mean ± s.e.m.); $n = 25$) neurons tested, capsaicin produced a prolonged (10 ± 3 s) inward current (Fig. 1a, upper panel), verifying the expression of TRPV1 receptors. Bath application of

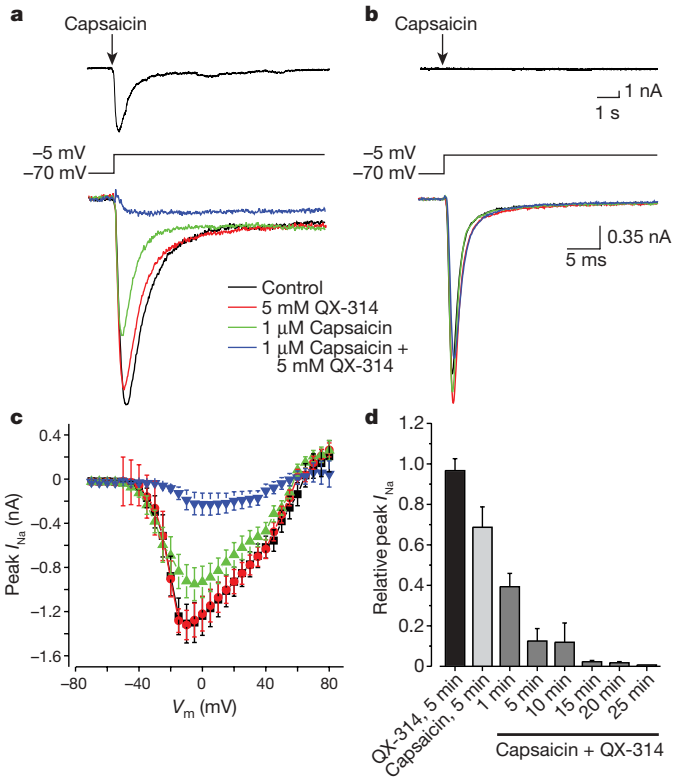


Figure 1 | Application of extracellular QX-314 (5 mM) and capsaicin (1 μM) together selectively blocks sodium currents in capsaicin-responsive DRG neurons. a, Lower panel: effect on sodium current (elicited by a step from −70 to −5 mV) of a 10-min wash-in of 5 mM QX-314 alone (red trace), 1 μM capsaicin alone (green trace), and 5mM QX-314 and 1 μM capsaicin applied together (blue trace) in a small (24 μm) capsaicin-sensitive cultured adult DRG neuron. Black trace, control. Upper panel: capsaicin induced a prolonged inward current (holding voltage −70 mV) in this neuron. **b**, Lack of effect on sodium current of the same drug applications on a large (52 μm) capsaicin-insensitive neuron. **c**, Peak inward current as a function of test pulse recorded in control (black), in the presence of 5 mM QX-314 alone (red; 5 min application), 1 μM capsaicin alone (green; 5 min application), and 5 mM QX-314 and 1 μM capsaicin applied together (blue; 5 min application). Results are means ± s.e.m. for 25 small capsaicin-sensitive neurons. Currents were elicited by 150-ms depolarizing steps from a holding potential of −70 mV to a range of test potentials in 5-mV increments. **d**, Time course of the combined effect of capsaicin and QX-314 on peak sodium current. Results are means and s.e.m. of peak sodium current relative to control. ($n = 25$).

[1]Department of Anesthesia and Critical Care, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts 02129, USA. [2]Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA.

©2007 Nature Publishing Group

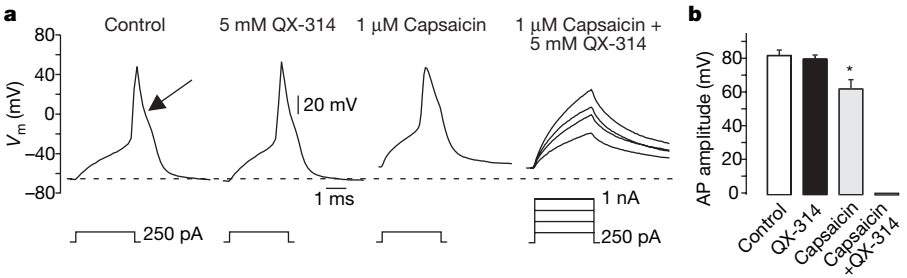


Figure 2 | Application of QX-314 and capsaicin together blocks excitability in nociceptor neurons. a, First panel: a depolarizing current step (250 pA, 4 ms) applied to a small (23 μm) DRG neuron evoked a nociceptor-like broad action potential with a prominent deflection on the falling phase (arrow). Second panel: application of QX-314 (5 mM, applied for 2 min) had no effect. Third panel: capsaicin (1 μM, 5 min) decreased the action potential amplitude, probably as a result of a modest decrease in sodium current and inactivation of sodium current secondary to capsaicin-induced depolarization. Fourth panel: QX-314 and capsaicin (2 min) applied together completely abolished action potential generation even with larger current injections. **b**, Mean action potential amplitudes. Results are means and s.e.m. ($n = 25$ for QX-314; $n = 15$ for capsaicin and capsaicin plus QX-314). Asterisk, $P < 0.05$ (Student's t-test).

5 mM QX-314 alone had little effect on sodium current elicited by a depolarizing step (decrease by 3% ± 0.5% after application for 5 min; $n = 25$) (Fig. 1a, c). Application of capsaicin alone (1 μM for 5 min) decreased sodium current moderately (31% ± 9% inhibition ($n = 25$), which is consistent with earlier data from nociceptors[9]. However, when QX-314 was applied together with capsaicin, the sodium current was almost totally blocked (inhibition by 98% ± 0.4%; $n = 25$) (Fig. 1a, c). As expected if the block of sodium current resulted from gradual entry of QX-314 through TRPV1 receptors, inhibition developed over several minutes and was nearly complete after 15 min (Fig. 1d).

In large DRG neurons (with a soma diameter of more than 40 μm) (Fig. 1b), capsaicin neither elicited an inward current (10 of 10) nor altered sodium current (3% ± 2% increase after application for 10 min; $n = 10$), as reported[9]. Application of QX-314 and capsaicin together had little or no effect on sodium current in large-diameter neurons (decrease by 9% ± 5% after application for 10 min; $n = 10$). The lack of effect reflects the failure of QX-314 to enter the large neurons, because a fivefold lower concentration of QX-314 (1 mM) applied intracellularly by inclusion in the recording pipette produced a complete block of sodium current within 30–60 s (10 of 10 cells; 48 ± 7 μm diameter; data not shown). Thus, the ability of QX-314 and capsaicin applied together to inhibit sodium current is restricted to neurons expressing TRPV1 receptors, as expected if QX-314 enters neurons through TRPV1 receptors. In current-clamp experiments using physiological internal and external solutions, QX-314 and capsaicin applied together blocked action potential generation completely in small-diameter neurons (Fig. 2; 15 of 15 neurons).

We next examined whether the combination of capsaicin and QX-314 can reduce pain behaviour *in vivo*. Injection of QX-314 alone (10 μl of 2% solution) into the hindpaw of adult rats had no significant effect on the mechanical threshold for eliciting a withdrawal response, as determined by von Frey hairs ($P = 0.33$; Fig. 3a). Capsaicin alone (1 μg μl^{-1}) elicited spontaneous flinching (40 ± 6 flinches in 5 min), reflecting the direct irritant action of the capsaicin on nociceptors, and after 15 and 30 min it significantly decreased the mechanical threshold ($P < 0.05$; Fig. 3a), as reported previously[10]. Injection of capsaicin and QX-314 together did not significantly change the number of flinches during the first 5 min after the injection (30 ± 7, $P = 0.24$). However, the combination completely abolished the later decrease in mechanical threshold normally produced by capsaicin alone ($P = 0.14$, measured at 15 min). Moreover, 60 min after the combined injection of capsaicin and QX-314 the mechanical threshold actually increased, reaching twice the baseline value 2 h after injection (46 ± 5 g versus 24 ± 3 g, $P < 0.05$). In three animals the paw was insensitive to even the highest-value von Frey filament (57 g). The elevated mechanical threshold lasted for about 3 h and then gradually returned to basal levels by 4 h (Fig. 3a).

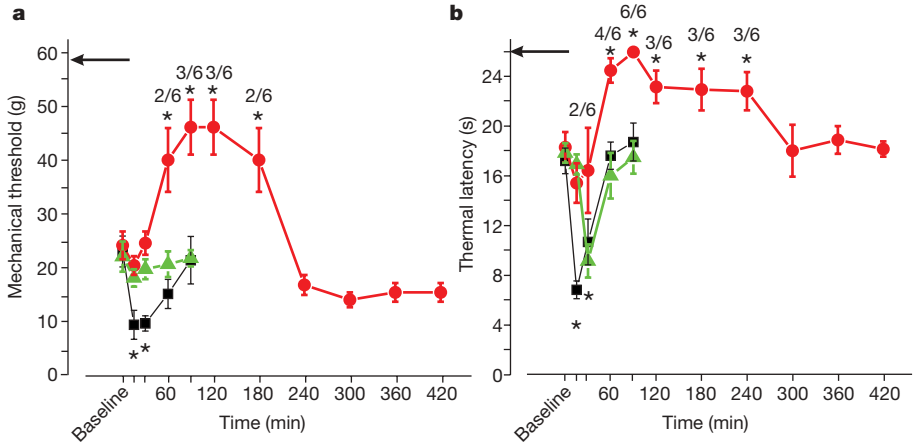


Figure 3 | Intraplantar injection of capsaicin and QX-314 together produces prolonged local anaesthesia to mechanical (von Frey) and thermal noxious stimuli. a, Mechanical threshold for paw withdrawal after intraplantar injection of QX-314 alone (2%, 10 μl; green), capsaicin alone (1 μg μl^{-1}; black), or QX-314 and capsaicin applied together (red). Results are means ± s.e.m. ($n = 6$ for each group; asterisk, $P < 0.05$ (ANOVA followed by Dunnett's test)). The numbers of animals that did not respond to the highest-value von Frey filament (57 g, arrow) are indicated. **b**, Latency for paw withdrawal from radiant heat (controlled radiant heat stimulus focused on an 8 mm × 8 mm spot on the plantar surface). Results are means ± s.e.m. ($n = 6$ for each group; asterisk, $P < 0.05$ (ANOVA followed by Dunnett's test)). The arrow indicates a cutoff of 26 s; the numbers of animals not responding within this time are indicated.

©2007 Nature Publishing Group

Similar effects were found for a controlled radiant noxious heat stimulus. QX-314 alone transiently reduced the thermal response latency 30 min after the injection ($P < 0.01$ at 30 min; $P > 0.05$ for all other time points; Fig. 3b). Capsaicin (1 μg μl^{-1}) alone also, as expected, decreased the thermal response latency ($P < 0.01$ at 15 and 30 min; Fig. 3b). However, whereas both QX-314 and capsaicin administered alone increased sensitivity to heat, the application of QX-314 and capsaicin together progressively anaesthetized the animals to noxious heat such that 2 h after the injection no animal reacted to the radiant noxious heat stimulus applied for 25 s. This effect remained for 4 h after injection (Fig. 3b).

We next tested whether the administration of capsaicin and QX-314 together could be used to produce a regional nerve block without the motor deficits produced by lidocaine. Injection of standard 2% lidocaine solution in close proximity to the sciatic nerve caused complete paralysis of the lower limb at 15 min (six of six animals) and complete or partial paralysis was still present at 30 min (motor score 1.67 ± 0.2, $P < 0.01$; Fig. 4c). There was a complete loss of the tactile-stimulus-evoked placing reflex for at least 30 min in all animals, with full recovery of the sensory and motor deficits by 45 min (Fig. 4). In pilot experiments it became clear that although QX-314 is a simple derivative of lidocaine with a similar if not higher 50% inhibitory concentration (IC_{50}) for sodium channels, much lower concentrations of QX-314 could be used to produce effective local anaesthesia when applied with capsaicin. Injection of QX-314 (0.2%, 100 μl) alone had no effect on motor function (six of six animals; Fig. 4c) or on mechanical threshold ($P = 0.7$) and noxious thermal response latency ($P = 0.66$; Fig. 4 a, b). Capsaicin alone (0.5 μg μl^{-1}, 100 μl) injected near the nerve decreased both the mechanical threshold ($P < 0.05$) and the thermal latency ($P < 0.05$) for 30 min after injection (Fig. 4a, b). During this period four of the six animals showed a sustained flexion of the injected limb leading to a slight impairment of locomotion (mean motor score 0.7 ± 0.2, $P < 0.01$), but movement of the knee and hip as well as the placing reflex were unchanged. We interpret the sensitivity and motor changes as reflecting the activation of nociceptor axons producing a sustained flexion reflex. For the application of QX-314 and capsaicin together into the parasciatic nerve region, we injected QX-314 first, followed 10 min later by capsaicin, with the idea that QX-314 would be present extracellularly and ready to enter TRPV1 channels as soon as they were activated. Indeed, there was little or no behavioural response to the capsaicin injection when preceded by QX-314 injection, and an effective anaesthesia to noxious stimuli developed with a very marked increase in mechanical threshold (all animals showed no response to the stiffest von Frey hair (57 g) compared with a pre-injection withdrawal threshold of 15.2 ± 3.4 g; $P < 0.01$, $n = 6$), and in thermal response latency (22.3 ± 2.3 s versus 14.9 ± 0.4 s; $P < 0.05$, $n = 6$). The changes were evident 15 min after the capsaicin injection for mechanical stimuli and at 30 min for thermal stimuli, and lasted for 90 min (Fig. 4a, b). Five of six animals had no motor deficit (mean motor score 0. 17 ± 0.17, $P = 0.34$; Fig. 4c) or a change in the placing reflex. One animal showed sustained flexion similar to, but more transient than, that observed when capsaicin was injected alone.

These results show that excitability of nociceptors can be selectively inhibited by a membrane-impermeant local anaesthetic derivative when it is applied together with an agonist of TRPV1 receptors, and that this effect is restricted to neurons expressing TRPV1 receptors. Because QX-314 has no effect on neuronal sodium channels when present externally but blocks from the inside of the membrane[4–8], it seems clear that the application of QX-314 together with capsaicin results in a rapid entry of QX-314 molecules into the neurons. The simplest possibility is that QX-314 permeates directly through TRPV1 channels opened by capsaicin. This seems plausible because the pore of TRPV1 channels is unusually wide compared

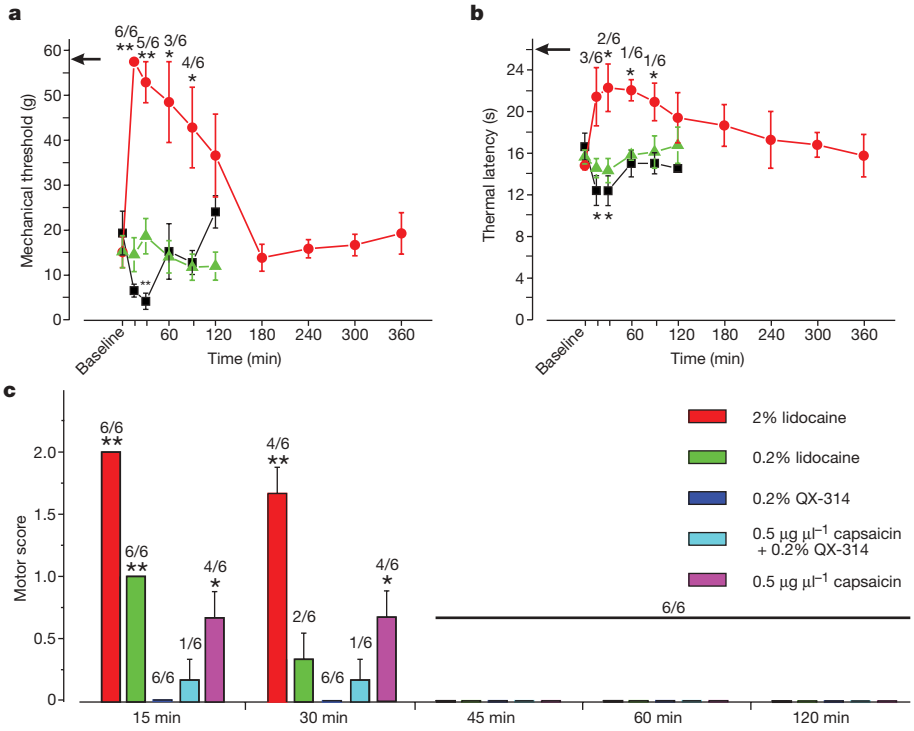


Figure 4 | Injection of QX-314 followed by capsaicin adjacent to the sciatic nerve anaesthetized the animals to noxious mechanical and thermal stimuli of the hindlimbs without producing any motor deficit. a, Mechanical threshold for paw withdrawal after sciatic injection of QX-314 alone (0.2%, 100 μl; green), capsaicin alone (0.5 μg μl^{-1}, 100 μl; black) or QX-314 injected 10 min before capsaicin (red). Results are means ± s.e.m. ($n = 6$ for each group; asterisk, $P < 0.05$; two asterisks, $P < 0.01$ (ANOVA followed by Dunnett's test)). The numbers of animals that did not respond to the highest force (57 g, arrow) are indicated. **b**, Latency for paw withdrawal from radiant heat. Symbols are as in **a**. **c**, Change in motor function (score: 2, full paralysis; 1, partial paralysis; 0, no impairment) evaluated after sciatic injection of lidocaine (2% or 0.2%), QX-314 (0.2%), capsaicin (0. 5 μg μl^{-1}), or QX-314 followed by capsaicin. Results are means and s.e.m. ($n = 6$ for each group; asterisk, $P < 0.05$; two asterisks, $P < 0.01$ (ANOVA followed by Dunnett's test, referenced to animals before injection; that is, motor score of 0)). Numbers of animals affected are indicated above columns.

©2007 Nature Publishing Group

with other cation-selective channels (including other members of the TRP family)[11] and is large enough to allow permeation of the large cations tetraethylammonium and *N*-methyl-D-glucamine[12]. QX-314 can permeate slowly through the pore of the cardiac ($Na_v1.5$) sodium channel[8,13], which probably has a smaller pore than TRPV1 channels. Direct entry of QX-314 would also be consistent with the proposal that that the dye FM1-43, with a larger molecular mass than QX-314, permeates through the pore of TRPV1 channels[3].

Our results show that the application of QX-314 and capsaicin together produces long-lasting decreases in the response to painful mechanical and thermal stimuli, and the regional anaesthesia produced by this mechanism is without the motor deficits that accompany conventional local anaesthesia with lidocaine and other lipid-soluble local anaesthetics. This strategy for blocking pain should be advantageous for generating pain-restricted local anaesthesia when preserving motor and autonomic responses and non-painful sensations is desirable, such as in childbirth and some dental procedures, as well as in treating nociceptor-driven chronic pain such as postherpetic neuralgia.

METHODS SUMMARY

Electrophysiology. Whole-cell voltage-clamp or current-clamp recordings were made at room temperature (21–23 °C) from primary cultures of neurons from DRGs from 6–8-week-old Sprague–Dawley rats. Voltage-clamp recordings used solutions designed to isolate sodium currents by blocking potassium and calcium currents and with decreased external sodium concentrations to improve voltage clamp. Current-clamp recordings were made by using solutions with physiological ionic composition. Solution changes were made in less than 1 s with a multibarrel drug delivery system.

Behaviour. For intraplantar injections, rats were habituated to handling and tests were performed with the experimenter blind to the treatment. Intraplantar injections of vehicle (20% ethanol, 5% Tween 20 in saline; 10 μl), capsaicin (1 μg μl^{-1}), QX-314 (2%) or a mixture of capsaicin and QX-314 were made into the left hindpaw. The mechanical threshold was determined with von Frey filaments[14] and thermal sensitivity was tested by latency to paw withdrawal from a controlled radiant heat source[15] focused on an 8 mm × 8 mm spot on the plantar surface.

For sciatic nerve injections, animals were habituated to handling for 10 days, and lidocaine (0.2% or 2%; 100 μl), QX-314 (0.2%, 100 μl) alone, capsaicin (50 μg in 100 μl) alone, or QX-314 followed by capsaicin (10-min interval) were injected into the area of sciatic nerve below the hip joint. Mechanical threshold was determined with von Frey filaments, and thermal sensitivity was tested by latency to paw withdrawal from a radiant heat source. Motor function of the injected leg was assessed every 15 min with a grading score of 0 (no effect; normal gait and limb placement), 1 (limb movement but with abnormal placement and movement) or 2 (complete loss of limb movement)[16]. Walking, climbing and the placing reflex were examined.

Statistical analysis. Data are presented as means ± s.e.m.. Statistics were analysed with Student's *t*-test or, when appropriate, one-way analysis of variance followed by Dunnett's test.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 June; accepted 28 August 2007.

1. Hille, B. Local anesthetics: hydrophilic and hydrophobic pathways for the drug-receptor reaction. *J. Gen. Physiol.* **69,** 497–515 (1977).
2. Caterina, M. J. & Julius, D. The vanilloid receptor: a molecular gateway to the pain pathway. *Annu. Rev. Neurosci.* **24,** 487–517 (2001).
3. Meyers, J. R. *et al.* Lighting up the senses: FM1-43 loading of sensory cells through nonselective ion channels. *J. Neurosci.* **23,** 4054–4065 (2003).
4. Frazier, D. T., Narahashi, T. & Yamada, M. The site of action and active form of local anesthetics. II. Experiments with quaternary compounds. *J. Pharmacol. Exp. Ther.* **171,** 45–51 (1970).
5. Strichartz, G. R. The inhibition of sodium currents in myelinated nerve by quaternary derivatives of lidocaine. *J. Gen. Physiol.* **62,** 37–57 (1973).
6. Yeh, J. Z. Sodium inactivation mechanism modulates QX-314 block of sodium channels in squid axons. *Biophys. J.* **24,** 569–574 (1978).
7. Cahalan, M. D. & Almers, W. Interactions between quaternary lidocaine, the sodium channel gates, and tetrodotoxin. *Biophys. J.* **27,** 39–55 (1979).
8. Qu, Y., Rogers, J., Tanada, T., Scheuer, T. & Catterall, W. A. Molecular determinants of drug access to the receptor site for antiarrhythmic drugs in the cardiac Na^+ channel. *Proc. Natl Acad. Sci. USA* **92,** 11839–11843 (1995).
9. Liu, L., Oortgiesen, M., Li, L. & Simon, S. A. Capsaicin inhibits activation of voltage-gated sodium currents in capsaicin-sensitive trigeminal ganglion neurons. *J. Neurophysiol.* **85,** 745–758 (2001).
10. Sluka, K. A. & Willis, W. D. The effects of G-protein and protein kinase inhibitors on the behavioral responses of rats to intradermal injection of capsaicin. *Pain* **71,** 165–178 (1997).
11. Owsianik, G., Talavera, K., Voets, T. & Nilius, B. Permeation and selectivity of TRP channels. *Annu. Rev. Physiol.* **68,** 685–717 (2006).
12. Hellwig, N. *et al.* TRPV1 acts as proton channel to induce acidification in nociceptive neurons. *J. Biol. Chem.* **279,** 34553–34561 (2004).
13. Sunami, A., Glaaser, I. W. & Fozzard, H. A. A critical residue for isoform difference in tetrodotoxin affinity is a molecular determinant of the external access path for local anesthetics in the cardiac sodium channel. *Proc. Natl Acad. Sci. USA* **97,** 2326–2331 (2000).
14. Amaya, F. *et al.* The voltage-gated sodium channel $Na_v1.9$ is an effector of peripheral inflammatory pain hypersensitivity. *J. Neurosci.* **26,** 12852–12860 (2006).
15. Hargreaves, K., Dubner, R., Brown, F., Flores, C. & Joris, J. A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. *Pain* **32,** 77–88 (1988).
16. Hara, K., Saito, Y., Kirihara, Y. & Sakura, S. The interaction between γ-aminobutyric acid agonists and diltiazem in visceral antinociception in rats. *Anesth. Analg.* **98,** 1380–1384 (2004).

Acknowledgements We thank T. Herbert and S. Lin for technical assistance.

Author Contributions B.P.B. conceived of introducing charged ion channel blockers through TRPV1 channels; A.M.B., C.J.W. and B.P.B. designed the experiments; A.M.B. did all experiments; and A.M.B., B.P.B. and C.J.W. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to B.P.B. (bruce_bean@hms.harvard.edu).

©2007 Nature Publishing Group

METHODS

Electrophysiology. DRGs from 6–8-week-old Sprague–Dawley rats were removed and placed in DMEM containing 1% penicillin–streptomycin (Sigma), treated for 90 min with 5 mg ml^{-1} collagenase, 1 mg ml^{-1} Dispase II (Roche), then with 0.25% trypsin for 7 min, followed by 2.5% trypsin inhibitor. Cells were triturated in the presence of DNAase I inhibitor (50 U), centrifuged through 15% BSA (Sigma), resuspended in 1 ml of Neurobasal medium (Sigma), 10 μM AraC, 50 ng ml^{-1} nerve growth factor and 2 ng ml^{-1} GDNF (glial cell line-derived neurotrophic factor) and plated onto 35-mm tissue culture dishes (Becton Dickinson) coated with 500 μg ml^{-1} polylysine and 5 mg ml^{-1} laminin, at 8,000–9,000 cells per dish. Cultures were incubated at 37 °C under 5% CO_2. Recordings were made at room temperature within 48 h of plating.

Whole-cell voltage-clamp or current-clamp recordings were made with an Axopatch 200A amplifier (Molecular Devices) and patch pipettes with resistances of 1–2 MΩ. For voltage-clamp recordings the pipette capacitance was decreased by wrapping the shank with Parafilm or coating the shank with Sylgard (Dow Corning). Cell capacitance was compensated for with the use of the amplifier circuitry, and linear leakage currents were subtracted with a *P*/4 procedure. Series resistance (usually 3–7 MΩ and always less than 10 MΩ) was compensated for by about 80%. Voltage-clamp recordings used solutions designed to isolate sodium currents by blocking potassium and calcium currents and with a decreased external sodium concentration, to improve voltage clamp. Pipette solution was 110 mM CsCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 11 mM EGTA, 10 mM HEPES, pH adjusted to 7.4 with about 25 mM CsOH. The external solution was 60 mM NaCl, 60 mM choline chloride, 4 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 0.1 mM $CdCl_2$, 15 mM tetraethylammonium chloride, 5 mM 4-aminopyridine, 10 mM glucose, 10 mM HEPES, pH adjusted to 7.4 with NaOH. No correction was made for liquid junction potential (−2.2 mV).

Current-clamp recordings were made with the fast current-clamp mode of the Axopatch 200A amplifier by using a pipette solution of 135 mM potassium gluconate, 2 mM $MgCl_2$, 6 mM KCl, 10 mM HEPES, 5 mM MgATP, 0.5 mM Li_2GTP, pH adjusted to 7.4 with KOH, and an external solution of 145 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM HEPES, 10 mM glucose, pH adjusted to 7.4 with NaOH. Membrane potential was corrected for the liquid junction potential (−15 mV).

Command protocols were generated and data were digitized with a Digidata 1200 A/D interface with pCLAMP 8.2 software (Molecular Devices). Voltage-clamp current records were low-pass filtered at 2 kHz, and current-clamp recordings were low-pass filtered at 10 kHz (−3 dB, four-pole Bessel filter).

Drugs were applied with a multibarrel drug delivery system placed 200–250 μm from the neuron. Solution exchange was complete in less than 1 s.

Behaviour. For intraplantar injections, thermal sensitivity was determined with the Hargreaves test method[15] (Ugo Basile Biological Research Apparatus).

For sciatic nerve injections, it was not possible to assay sensory sensitivity during the period of paralysis produced by lidocaine.

©2007 Nature Publishing Group

LETTERS

Snapshots of nuclear pore complexes in action captured by cryo-electron tomography

Martin Beck[1], Vladan Lučić[1], Friedrich Förster[1], Wolfgang Baumeister[1] & Ohad Medalia[1]†

Nuclear pore complexes reside in the nuclear envelope of eukaryotic cells and mediate the nucleocytoplasmic exchange of macromolecules[1]. Traffic is regulated by mobile transport receptors that target their cargo to the central translocation channel, where phenylalanine-glycine-rich repeats serve as binding sites[2]. The structural analysis of the nuclear pore is a formidable challenge given its size, its location in a membranous environment and its dynamic nature. Here we have used cryo-electron tomography[3] to study the structure of nuclear pore complexes in their functional environment, that is, in intact nuclei of *Dictyostelium discoideum*. A new image-processing strategy compensating for deviations of the asymmetric units (protomers) from a perfect eight-fold symmetry enabled us to refine the structure and to identify new features. Furthermore, the superposition of a large number of tomograms taken in the presence of cargo, which was rendered visible by gold nanoparticles, has yielded a map outlining the trajectories of import cargo. Finally, we have performed single-molecule Monte Carlo simulations of nuclear import to interpret the experimentally observed cargo distribution in the light of existing models for nuclear import.

A variety of electron microscopy approaches has been used to elucidate the overall architecture of nuclear pore complexes (NPCs)[1] but substantial improvements in resolution are needed to use three-dimensional electron microscopy maps as a frame into which high-resolution structures of individual nucleoporins or of larger modules can be fitted with confidence. Hitherto, most electron microscopy studies used preparations derived from isolated nuclear envelopes, often by detergent extraction, which bears the risk of artefacts such as the loss of components, and it compromises our ability to perform studies under physiologically relevant conditions. Cryo-electron tomography is an imaging modality that combines the potential of three-dimensional imaging with a close-to-life preservation of biological samples, and it allows the investigation of large structures with unique topologies (such as organelles or even whole cells) at molecular resolution[3]. We have shown previously that cryo-electron tomography can be applied to intact, transport-competent nuclei from *Dictyostelium discoideum*, providing new insights into the structure and dynamics of the NPC[4].

It had been noted before that NPCs exhibit plasticity[5] and show conspicuous deviations from eight-fold symmetry to the extent that supernumerary protomers are incorporated into the complex[6]. To avoid the fact that deviations from rotational symmetry limit resolution, we used an image-processing strategy akin to the correlation averaging or unbending of distorted lattices[7]. An initial reference was created by averaging tomograms of 523 NPCs with an imposed eight-fold symmetry; from this average, a segment was cut out comprising one complete protomer flanked by two half protomers. This reference was then used in a refinement cycle. After mapping 4,184 asymmetric units individually and aligning them, a new average was created. The improvement in resolution that can be achieved depends on the accuracy with which individual protomers can be located and on their mode of accommodating distortions (non-rigid compared with rigid body movements)[8]. The significant improvement in resolution that we obtained (from 8.3 nm to 5.8 nm; Supplementary Fig. 3) indicates that individual protomers can indeed be located with sufficient accuracy. The measured displacement of the protomers (4.6 nm mean in-plane displacement from symmetry-related positions) is consistent with these values (Fig. 1a, b). Out-of-plane displacements of the protomers seem to be negligible. Most NPCs assume a slightly elliptical shape that entails a small change of the pore diameter as compared to a perfectly symmetric ring structure (Fig. 1c, d). Notably, a recent structural study of Nup58/45 indicates that the diameter of the central transport channel is adjusted by circumferential sliding of this subcomplex[9]. The refined component averages were reassembled computationally into a complete model of the NPC (see below); in an animated representation, the measured distortions were translated into motions (see Supplementary Movie 1).

With the improvement in resolution, several new features were revealed: the cytoplasmic ring, spoke ring and nuclear ring are now seen as distinct but nevertheless connected structures (Fig. 2a, b). In contrast to the nuclear ring, the cytoplasmic ring seems to be only weakly connected to the spoke ring. The spokes form a ring with an inner diameter of 75 nm, whereas the fused inner nuclear membrane and outer nuclear membrane form a ring of 105 nm in diameter. A cut-away view of a protomer (Fig. 2c) shows the interactions of the three rings with the membrane. Each spoke has a clamp-like shape and is attached to the membrane at two specific sites. The spoke ring apparently stabilizes the membrane curvature, a task previously ascribed to the Nup107 subcomplex[10]. Interestingly, Nup133, a member of this subcomplex, contains an amphipathic alpha-helical motif for sensing membrane curvature[11]. The contact region between the membrane and the spoke ring is a porous, sponge-like structure. In a three-dimensional surface-rendered view, it seems to be traversed by a channel of approximately 9 nm in diameter (Fig. 2c and Supplementary Fig. 2a). It has been suggested previously that an alternative passageway exists that is used exclusively by small molecules, allowing them to bypass the meshwork of phenylalanine-glycine repeats residing in the central channel[12]. We have not found evidence suggesting that classical import cargo can be transported through these regions (see below). However, recent observations[13] suggest that import of inner nuclear membrane proteins proceeds along a path near the NPC adjacent to the membrane.

A striking feature is the existence of a luminal connector element spanning the space between the inner nuclear membrane and outer nuclear membrane near the periphery of the NPC (Fig. 2c, e). It is

[1]Max Planck Institute of Biochemistry, D-82152 Martinsried, Germany. †Present address: The Department of Life Sciences, The Ben Gurion University and the NIBN, Beer Sheva 84105, Israel.

©2007 Nature Publishing Group

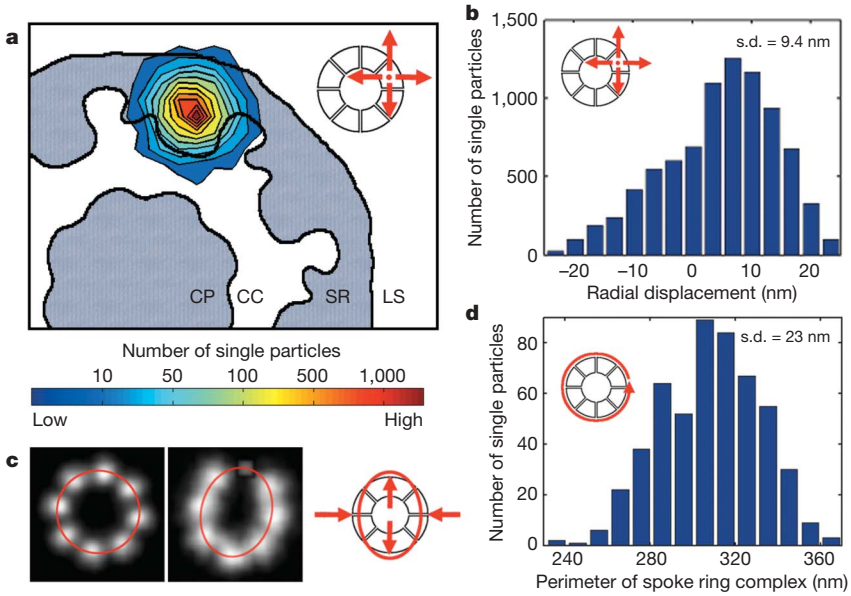


Figure 1 | Plasticity of intact NPCs is revealed by displacement analysis. **a**, Heat map showing the displacement of a NPC asymmetric unit (protomer) from the ideal eight-fold rotational symmetric arrangement within the spoke ring plane (superimposed). CC, central channel; CP, central plug; LS, luminal space; SR, spoke ring. **b**, Distribution of the radial displacement of the asymmetric units (mean 4.6 nm, s.d. 9.4 nm). **c**, Elliptical distortions of individual NPCs visualized as the probability density of the positions of the eight protomers (white). Left: when eight-fold rotational symmetry is imposed, the positions of the protomers form a circle (red). Right: protomer positions form an ellipse (red) when the alignment is performed with individual NPCs. The subunits located at one end of their major (long) axes were used as reference points for the alignment (see the top protomer on the figure). **d**, Distribution of the perimeter of the spoke ring complex of individual NPCs.

attached to the membrane at the opposite sides of the membrane contacts of the cytoplasmic and nuclear rings, correspondingly. In view of the otherwise frail contacts, this connector may be a key element in the stabilization of the whole structure. It might define the distance of 30 nm between the inner nuclear membrane and outer nuclear membrane and contribute to the maintenance of the

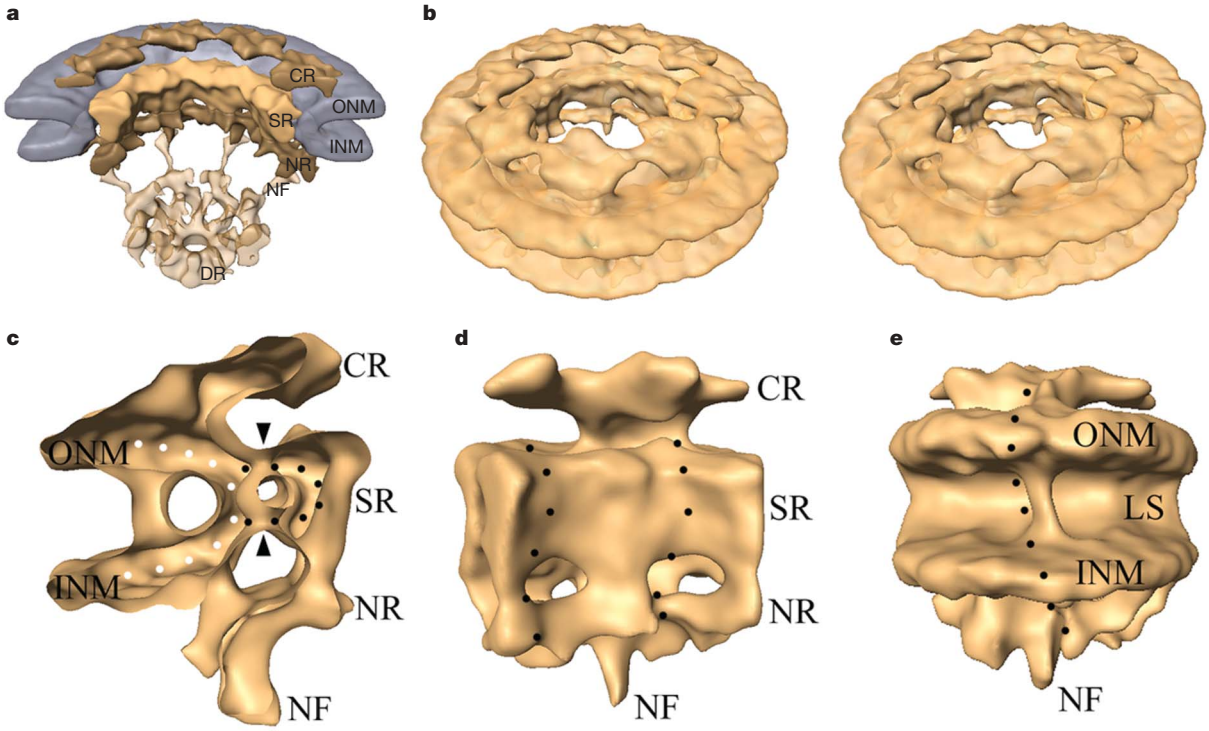


Figure 2 | Refined structure of the *Dictyostelium* NPC. **a**, Cut-away view of the structure of rejoined asymmetric units with subjective segmentation for the cytoplasmic (CR), spoke (SR) and nuclear rings (NR) in shades of yellow, and the inner (INM) and outer nuclear membrane (ONM) in bluish-grey. For clarity, the CP/T has been omitted and the basket with nuclear filaments (NF) and distal ring (DR) was rendered transparent. **b**, Stereo view of the scaffold structure of the NPC. **c**, Cut-away view of a protomer. The fused inner nuclear membrane and outer nuclear membrane and the clamp-shaped spoke structure are indicated by white dots and black dots, respectively; arrowheads mark the entry and exit of an apparent channel. **d**, Averaged structure of a protomer flanked by two half protomers seen from the central channel (black dots mark the protomer boundaries). **e**, View from the luminal space (LS). Black dots follow the rod-shaped structure that connects the cytoplasmic and nuclear rings.

©2007 Nature Publishing Group

membrane curvature. In contrast to earlier studies with *Xenopus* NPCs[5], the luminal structure observed with *Dictyostelium* NPCs is not ring-shaped but rod-shaped. The largest outer diameter of the entire NPC (135 nm) is defined by this membrane-spanning element.

Although tomography is normally a static tool, it can be used to take snapshots of processes, and, if these are properly superimposed, allows one to deduce the course of dynamic events: in this case, tomography has been used to visualize the trajectories of cargo as it traverses the NPC. To map the routes taken by cargo, we mapped the positions of a gold-labelled fusion protein consisting of a nuclear localization signal and two copies of green fluorescent protein (NLS–2×GFP) during translocation. This model substrate for the classical import pathway has been well characterized[14]. We determined the positions of 638 cargo molecules within 251 labelled NPCs and used the experimentally determined coordinates of individual gold-labelled cargoes to compute a three-dimensional probability density map of classical import cargo at the nuclear pore complex (Fig. 3 and Supplementary Fig. 4). The central plug or transporter (CP/T) is an electron-dense structure within the central channel that seems to be tripartite on the level of the cytoplasmic, spoke and nuclear rings, respectively[4]. The observed locations of cargo and the probability map derived from them match the shape of the CP/T electron microscopy density well, confirming that the mass observed within the central channel can at least partially be ascribed to cargo, as suggested previously[2]. However, the cargo density was relatively low in a region with a maximum diameter of 20 nm in the middle of the central channel (in plane with the spoke ring). Notably, the corresponding region of the CP/T exhibits a high electron microscopy density. It remains an open question whether this local mismatch was caused by displacement of NLS–2×GFP by other cargoes and/or nucleoporins[2]. Because our import reactions were carried out with cytosolic supernatant, other transport events involving (unlabelled) cargo, such as the export of ribosomal subunits and RNAs, are likely to proceed simultaneously with the import of the labelled cargo. In earlier studies, nucleoplasmin[15] and messenger ribonucleoproteins[16] caught in transport were both localized to the very centre of the central channel in individual NPCs. Three further observations can be derived from the cargo probability density. (1) The increased intensity at the cytoplasmic side in the area rich with cytoplasmic filaments suggests that these filaments supply initial docking sites for cargo complexes, as proposed earlier[17]. Although the cytoplasmic filaments are dispensable for nuclear import[18], they might increase the effective docking surface and possibly the residence time for import complexes at the cytoplasmic entry of the pore[19] and thereby speed up the (overall) translocation. (2) The gold-labelled cargoes we observed in the nuclear basket region were, in contrast to the cytoplasmic and central channel region, evenly distributed and without local enrichments, and therefore made only a minor contribution to the computed probability cloud. The lack of cargo density in the nuclear basket region can be

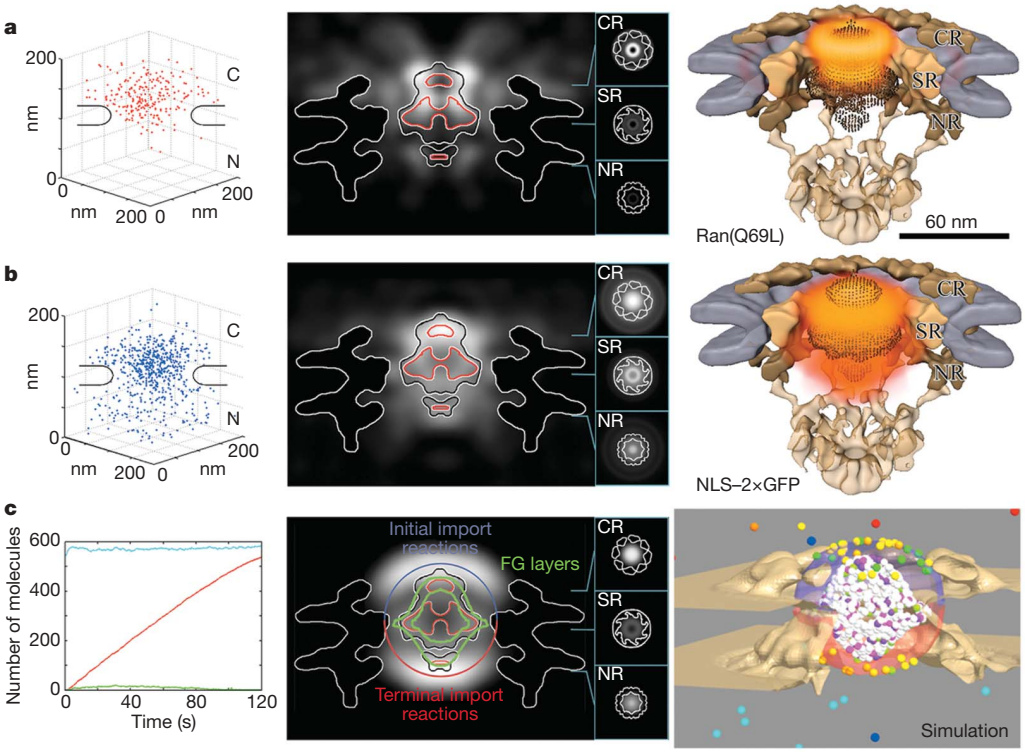


Figure 3 | Visualization of single molecules during nuclear import: experimental data and simulations. a, Gold-labelled Ran(Q69L)-GTP. Left: three-dimensional plot of the localization of Ran(Q69L)-GTP at the nuclear pore in the tomograms. Nuclear membranes are indicated on the sides. Middle: central slice along the nucleocytoplasmic axis of the (rotationally averaged) three-dimensional probability density map, superimposed onto the NPC structure. The three main electron-dense moieties of the central plug, in plane with the cytoplasmic (CR), spoke (SR) and nuclear rings (NR), are indicated by red lines and are shown as cross-sections in the insets. Right: three-dimensional probability distribution of Ran(Q69L)-GTP (orange clouds, brighter colour indicates higher probability) superimposed onto the central plug, which is shown as a dotted texture (brown). **b**, Gold-labelled classical import cargo NLS–2×GFP. Left, middle and right panels as for panel **a**. **c**, Computer simulations of nuclear import. Left: time courses of translocated substrate molecules (red), cargo (green) and RanGTP (cyan). Middle: central slice along the nucleocytoplasmic axis of the (rotationally averaged) three-dimensional probability density map of cargo, superimposed onto the NPC structure. Central plug as in the middle panels of **a** and **b**. In addition, cytoplasmic (blue), nuclear (red) and phenylalanine-glycine (FG) effector surfaces (green) are indicated. Right: surface-rendered cut-away view of a simulation 20 s after initialization. Small spheres represent individual molecules (or complexes). Substrate, red; cargo, green; importin-α, orange; importin-β, yellow; RanGDP, bright cyan; RanGTP, dark cyan; Cas-importin-α-RanGTP, pink; importin-β-RanGTP, purple; cargo complexes bound to phenylalanine-glycine sites, bright green; unoccupied phenylalanine-glycine sites, white. The NPC structure and membranes are shown in dark yellow, and the cytoplasmic (blue) and nuclear effector surfaces (red) are represented as half spheres.

©2007 Nature Publishing Group

explained either by a cargo release directly from the central channel, or by a mobile terminal cargo complex binding site[2]. (3) A gradient of binding sites with increasing affinity along the transport direction[20] would result in an increase in the probability density along the nucleocytoplasmic axis towards the nucleoplasm. We did not observe such a pattern in our experiments, although the intensity was higher at the exit of the central channel than within the central channel.

The GTP-hydrolysis of Ran at the cytoplasmic side of the NPC is the triggering event for the disassembly of export complexes[21]. In a control experiment (Fig. 3a), we mapped the positions of gold-conjugated Ran(Q69L)-GTP particles (127 labelled NPCs, 234 gold particles). This active site variant of Ran is unable to hydrolyse GTP and consequently can neither bind its import receptor NTF2 nor translocate into the nucleus in the conventional manner. It does interact with Ran-binding domains of the cytoplasmic filaments and is therefore expected to be localized mainly at the cytoplasmic side of the NPC[22]. The three-dimensional probability density of Ran(Q69L)-GTP indeed shows the highest intensities close to the outer part of the CP/T, in plane with the cytoplasmic ring. Nevertheless, we detected a small but significant number of gold-conjugated Ran(Q69L)-GTP particles inside the central channel, suggesting that they may enter the central channel by passive diffusion, albeit not very efficiently.

To clarify which parameters have a substantial impact on the local probability density of cargo, and illuminate some of the currently unresolved aspects of nuclear import, we used single-molecule three-dimensional Monte Carlo simulations of classical nuclear import in the microenvironment of an individual NPC (the shape of which was derived from the surface of our NPC structure). These simulations were done using the Mcell software package[23]. They tracked each molecule (or molecular complex) separately as it diffused through the pore and reacted with other molecules.

The three main nuclear import steps—the initial formation and docking of a cargo complex to the central channel, its interaction with the phenylalanine-glycine meshwork during the main translocation, and a terminal cargo release event[24]—were modelled in detail (Supplementary Table 1 and Supplementary Fig. 5). Essentially, there were only three parameters with undetermined values (see Supplementary Information). The values of other parameters—concentrations, diffusion constants, the permeabilities of the NPCs for diffusible molecules—as well as reaction constants (we modelled more than 50 reactions) were based on experimental data whenever possible. Some of them were adapted from earlier computer modelling[25–27] that did not take into account single molecules, the actual structure of the NPC, or the membrane to NPC surface ratio that ultimately determines the likelihood for a molecule to encounter the central channel. In our model, the relevant chemical reactions (Supplementary Fig. 5) are considered to occur at defined locations, namely the cargo complex formation and release at the entrance and exit of the central channel, and the interactions with the phenylalanine-glycine meshwork in between (effector surfaces shown in Fig. 3c). The output of our simulations provided the position of each molecule at each time point, and consequently, the three-dimensional distribution of cargo that was directly compared with the experimentally determined positions.

The agreement of the simulations with the experimental data turned out to depend critically on two parameter choices. First, as two NLS–2×GFP molecules are, on average, conjugated to each gold bead, one (monomolecular import) or two (bimolecular import) NLS sequences can be occupied by importin heterodimers during translocation. The bimolecular reaction is kinetically slower, causing a local increase of the probability density at the cytoplasmic face. Second, two different terminal cargo release rates are available in the literature: the (fast release) rate used in earlier simulations[27], and the experimentally determined (slow release) rate for Nup50 (ref. 28), which would cause a local increase of the probability density at the nuclear face.

Two out of six simulation scenarios produced probability densities that satisfactorily fitted the experimental data (see Supplementary Table 2 for a quantitative assessment). First, if the fast cargo release rate is used, a monomolecular cargo assembly reaction is necessary to obtain consistent intensities (Supplementary Fig. 6). Second, if a slow cargo release rate is assumed, at least partially bimolecular cargo assembly must be used (Fig. 3c). Whereas the latter case produces a better match (Supplementary Table 2), it is characterized by a slower flux rate (Supplementary Fig. 6), because at high substrate concentrations the importin concentration becomes rate limiting, resulting in a relatively slow initial event.

The selective phase model predicts that importins compensate for surface charges of substrates to allow the passage through a hydrophobic permeability barrier in the central channel, which is formed by a three-dimensional meshwork of phenylalanine-glycine-rich repeats with hydrogel-like properties[29]. Consequently, a high surface occupancy by importin-α/β heterodimers would result in a more effective transport of a cargo. Under conditions of excess substrate, one importin heterodimer may possibly tether a substrate to the phenylalanine-glycine meshwork, but the penetration of the meshwork would require the binding of a second importin heterodimer. The 'virtual gating' model[19] does not rely on a different chemical environment but on the affinity of cargo complexes to phenylalanine-glycine repeats. Therefore, the selective phase model is consistent with partially bimolecular import, whereas the virtual gating model is more consistent with monomolecular import. Recent findings suggest that both principles are involved in nuclear transport[30].

We have shown that compensating for deviations from perfect symmetry is necessary to increase the attainable resolution. Our analysis has led to the identification of novel structural features of the NPC. To understand fully the NPC architecture, high-resolution structures of nucleoporins or subcomplexes have to be fitted into the complete structure. Furthermore, we were able to visualize the classical import route in three dimensions and to provide clear experimental evidence suggesting that the electron optical density of the CP/T correlates with cargo. Classical import complexes are concentrated in the regions of the CP/T located in plane with the cytoplasmic and nuclear rings on entry into and release from the central channel. The ratios between the respective cargo probability intensities at the entrance, at the exit and within the central channel point to a slow initial event under conditions of excess substrate. For a further understanding of the transport mechanism and the nature of the CP/T, the three-dimensional localization of cargoes of other transport pathways, as well as transport factors and receptors, is needed.

METHODS SUMMARY

Dictyostelium discoideum cells were grown in AX2 medium and lysed by passing them through a 5 μm polycarbonate filter. Nuclei were enriched by centrifugation and, in the case of the cargo-tracing experiments, incubated in the presence of cytosolic supernatant and gold-labelled Ran(Q69L) or NLS–2×GFP, respectively. Afterwards the nuclei were pipetted onto lacy carbon grids and rapidly frozen in liquid ethane. Tilt series were collected with a Phillips CM 300 FEG or FEI Polara transmission electron microscope, typically covering an angular range from −63° to 63° and sampled with 1.5–2° tilt increments. Tomographic reconstructions were calculated by weighted back-projection and single-particle alignments were performed using 'missing-wedge'-weighted correlation averaging. During the structural refinement the eight asymmetric units of all single particles were aligned independently; resolution tests were performed by Fourier-shell correlation using the 0.5 criterion. Gold particles were localized in the aligned subvolumes and their positions were used to compute three-dimensional probability distributions of Ran(Q69L) or NLS–2×GFP at the nuclear pore. Single-molecule Monte Carlo simulations of classical nuclear import were done with the Mcell software. The cryo-electron microscopy map of the *Dictyostelium* nuclear pore served as a framework; diffusion coefficients, concentrations of the relevant molecules, kinetic constants and reaction schemes were based on published experimental data and computer simulations.

©2007 Nature Publishing Group

Detailed protocols, sample preparations, electron microscopy techniques, and image processing and simulation procedures are available in the Supplementary Information.

Received 4 June; accepted 17 August 2007.
Published online 12 September 2007.

1. Fahrenkrog, B., Koser, J. & Aebi, U. The nuclear pore complex: a jack of all trades. *Trends Biochem. Sci.* **29,** 175–182 (2004).
2. Fahrenkrog, B. & Aebi, U. The nuclear pore complex: nucleocytoplasmic transport and beyond. *Nature Rev. Mol. Cell Biol.* **4,** 757–766 (2003).
3. Lucic, V., Forster, F. & Baumeister, W. Structural studies by electron tomography: from cells to molecules. *Annu. Rev. Biochem.* **74,** 833–865 (2005).
4. Beck, M. *et al.* Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* **306,** 1387–1390 (2004).
5. Akey, C. W. Structural plasticity of the nuclear pore complex. *J. Mol. Biol.* **248,** 273–293 (1995).
6. Hinshaw, J. E. & Milligan, R. A. Nuclear pore complexes exceeding eightfold rotational symmetry. *J. Struct. Biol.* **141,** 259–268 (2003).
7. Saxton, W. O. & Baumeister, W. The correlation averaging of a regularly arranged bacterial cell envelope protein. *J. Microsc.* **127,** 127–138 (1982).
8. Saxton, W. O., Durr, R. & Baumeister, W. From lattice distortion to molecular distortion—characterizing and exploiting crystal deformation. *Ultramicroscopy* **46,** 287–306 (1992).
9. Melcak, I., Hoelz, A. & Blobel, G. Structure of Nup58/45 suggests flexible nuclear pore diameter by intermolecular sliding. *Science* **315,** 1729–1732 (2007).
10. Devos, D. *et al.* Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biol.* **2,** e380 (2004).
11. Drin, G. *et al.* A general amphipathic alpha-helical motif for sensing membrane curvature. *Nature Struct. Mol. Biol.* **14,** 138–146 (2007).
12. Stoffler, D. *et al.* Cryo-electron tomography provides novel insights into nuclear pore architecture: implications for nucleocytoplasmic transport. *J. Mol. Biol.* **328,** 119–130 (2003).
13. King, M. C., Lusk, C. P. & Blobel, G. Karyopherin-mediated import of integral inner nuclear membrane proteins. *Nature* **442,** 1003–1007 (2006).
14. Yang, W., Gelles, J. & Musser, S. M. Imaging of single-molecule translocation through nuclear pore complexes. *Proc. Natl Acad. Sci. USA* **101,** 12887–12892 (2004).
15. Rutherford, S. A., Goldberg, M. W. & Allen, T. D. Three-dimensional visualization of the route of protein import: the role of nuclear pore complex substructures. *Exp. Cell Res.* **232,** 146–160 (1997).
16. Kiseleva, E., Goldberg, M. W., Allen, T. D. & Akey, C. W. Active nuclear pore complexes in *Chironomus*: visualization of transporter configurations related to mRNP export. *J. Cell Sci.* **111,** 223–236 (1998).
17. Pante, N. & Aebi, U. Sequential binding of import ligands to distinct nucleopore regions during their nuclear import. *Science* **273,** 1729–1732 (1996).
18. Walther, T. C. *et al.* The cytoplasmic filaments of the nuclear pore complex are dispensable for selective nuclear protein import. *J. Cell Biol.* **158,** 63–77 (2002).
19. Rout, M. P., Aitchison, J. D., Magnasco, M. O. & Chait, B. T. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* **13,** 622–628 (2003).
20. Ben-Efraim, I. & Gerace, L. Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. *J. Cell Biol.* **152,** 411–417 (2001).
21. Matunis, M. J., Coutavas, E. & Blobel, G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J. Cell Biol.* **135,** 1457–1470 (1996).
22. Wu, J., Matunis, M. J., Kraemer, D., Blobel, G. & Coutavas, E. Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. *J. Biol. Chem.* **270,** 14209–14213 (1995).
23. Coggan, J. S. *et al.* Evidence for ectopic neurotransmission at a neuronal synapse. *Science* **309,** 446–451 (2005).
24. Becskei, A. & Mattaj, I. W. Quantitative models of nuclear transport. *Curr. Opin. Cell Biol.* **17,** 27–34 (2005).
25. Smith, A. E., Slepchenko, B. M., Schaff, J. C., Loew, L. M. & Macara, I. G. Systems analysis of Ran transport. *Science* **295,** 488–491 (2002).
26. Gorlich, D., Seewald, M. J. & Ribbeck, K. Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *EMBO J.* **22,** 1088–1100 (2003).
27. Riddick, G. & Macara, I. G. A systems analysis of importin-α-β mediated nuclear protein import. *J. Cell Biol.* **168,** 1027–1038 (2005).
28. Matsuura, Y. & Stewart, M. Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. *EMBO J.* **24,** 3681–3689 (2005).
29. Frey, S., Richter, R. P. & Gorlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314,** 815–817 (2006).
30. Patel, S. S., Belmont, B. J., Sante, J. M. & Rexach, M. F. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* **129,** 83–96 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Melchior for the Ran(Q69L) protein, S. Musser for the NLS–2×GFP plasmid, G. Gerisch and J. Glavy for valuable discussions, and A. Leis for critical reading of the manuscript. This work was supported in part by the European Union 3DEM Network of Excellence.

Author Information The structure has been deposited at the Macromolecular Structure database (EBI) under accession code EMD-1394. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to W.B. (baumeist@biochem.mpg.de) or O.M. (omedalia@bgu.ac.il).

©2007 Nature Publishing Group

LETTERS

Crystal structure of T4 endonuclease VII resolving a Holliday junction

Christian Biertümpfel[1], Wei Yang[1] & Dietrich Suck[2]

Holliday proposed a four-way DNA junction as an intermediate in homologous recombination[1], and such Holliday junctions have since been identified as a central component in DNA recombination and repair[2]. Phage T4 endonuclease VII (endo VII) was the first enzyme shown to resolve Holliday junctions into duplex DNAs by introducing symmetrical nicks in equivalent strands[3]. Several Holliday junction resolvases have since been characterized[4], but an atomic structure of a resolvase complex with a Holliday junction remained elusive. Here we report the crystal structure of an inactive T4 endo VII(N62D) complexed with an immobile four-way junction with alternating arm lengths of 10 and 14 base pairs. The junction is a hybrid of the conventional square-planar and stacked-X conformation. Endo VII protrudes into the junction point from the minor groove side, opening it to a 14 Å × 32 Å parallelogram. This interaction interrupts the coaxial stacking, yet every base pair surrounding the junction remains intact. Additional interactions involve the positively charged protein and DNA phosphate backbones. Each scissile phosphate that is two base pairs from the crossover interacts with a Mg^{2+} ion in the active site. The similar overall shape and surface charge potential of the Holliday junction resolvases endo VII, RuvC, Ydc2, Hjc and RecU, despite having different folds, active site composition and DNA sequence preference, suggest a conserved binding mode for Holliday junctions.

The 157-residue endo VII encoded by gene 49 of bacteriophage T4 resolves branched multimeric T4 DNA before packaging into phage heads[5,6]. Endo VII displays broad substrate specificity and can bind and cleave single-base mismatches in a DNA duplex as well as three-way and four-way branched DNA structures[7]. Crystal structures of endo VII alone revealed two elongated subunits, each consisting of six helices and a β-hairpin, intertwined into an S-shaped dimer[8]. The two subunits are antiparallel, with the amino terminus of one subunit

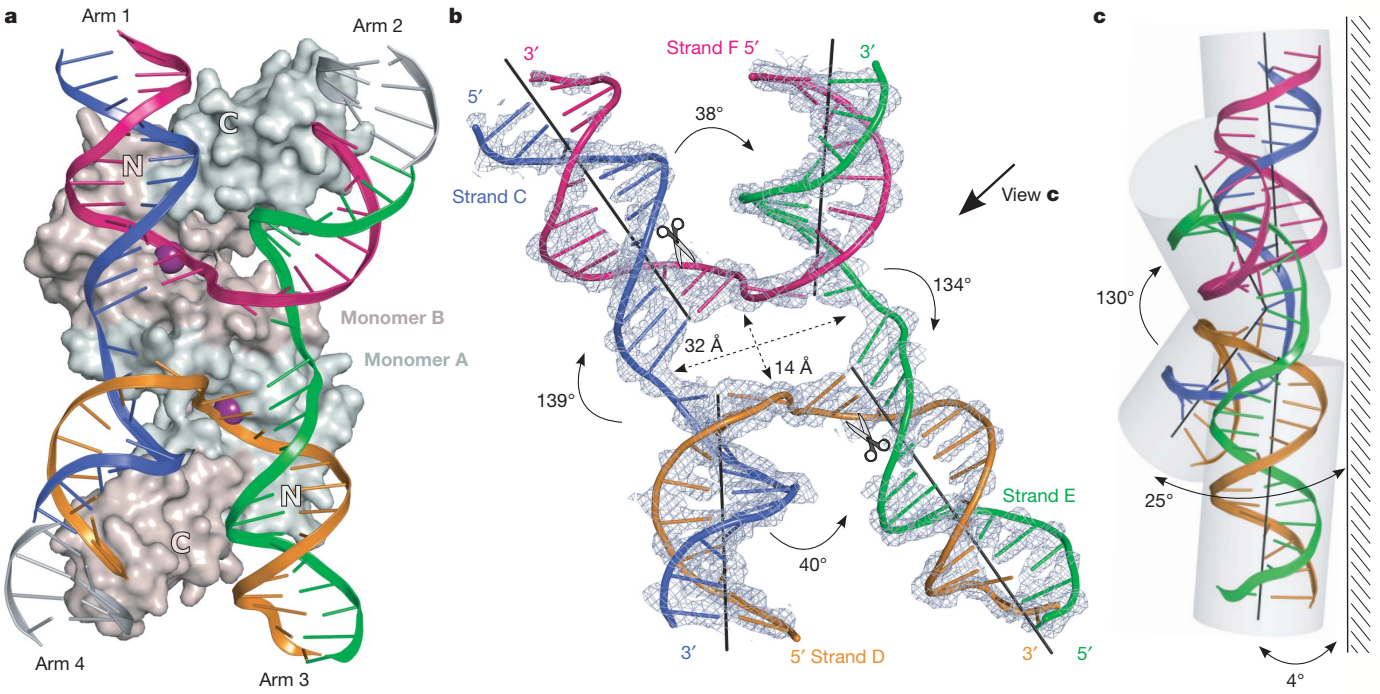


Figure 1 | Structure of the T4 endo VII(N62D)–Holliday junction complex. **a**, The front view of the complex shows the molecular surface of endo VII, with the two subunits coloured differently; the Holliday junction in ribbon diagrams, with each strand colour-coded. Symmetry-related DNAs are shown in grey, and Mg^{2+} ions are shown as purple spheres. **b**, The branch point of the junction is open. The angles between DNA arms are marked. Superimposed on the Holliday junction is the composite-omit electron density map contoured at 1.0 σ. Scissors indicate the scissile phosphates. **c**, A side view of the Holliday junction. Angles of DNA arms relative to a flat surface are shown.

[1]National Institute of Diabetes and Digestive and Kidney Diseases, Laboratory of Molecular Biology, 9000 Rockville Pike, Bethesda, Maryland 20892, USA. [2]European Molecular Biology Laboratory, Structural and Computational Biology Programme, Meyerhofstr. 1, 69117 Heidelberg, Germany.

©2007 Nature Publishing Group

interacting with the carboxy terminus of the other. In the absence of a DNA substrate, endo VII exhibits remarkable flexibility both in the centre and at the tips of the 'S'[9]. Holliday junctions in solution are also flexible[10,11]. In the absence of cations, Holliday junctions adopt an extended square-planar conformation. When cations are present to counteract charge repulsions, Holliday junctions exist in a co-axially stacked-X form, in which four arms are stacked in two pairs and each pair shares one helical axis.

To obtain substrate complexes, active-site residues of endo VII were individually mutated to make enzymes that were catalytically inactive but competent at binding Holliday junctions. By systematically varying the DNA arm lengths and pairing them with inactivated endo VII, we obtained 82 different forms of protein–DNA co-crystals. All but three diffracted X-rays poorly. The first crystal structure solved is the H43N endo VII bound to two DNA heteroduplexes, each composed of 14 base pairs (bp) plus two base mismatches (Supplementary Information and Supplementary Fig. 1). Using the protein part of this complex as a search model, we solved the structure of the endo VII(N62D) complexed with a Holliday junction that had arm lengths of alternating 10 bp and 14 bp. Each asymmetric unit contains one Holliday junction bound to a protein dimer (Fig. 1a), and the structure has been refined at 3.1 Å to $R_{work} = 24.5\%$ and $R_{free} = 27.2\%$ (see Supplementary Table 1). All residues of endo VII and most of the 48 bp DNA are traceable except for 1.5 bp at the end of a 14-bp arm (arm 1 in Fig. 1a).

The Holliday junction bound to endo VII is antiparallel, as defined by the polarities of the two non-exchanging strands (blue and green in Fig. 1). The phosphate backbones of the exchanging strands (red and orange in Fig. 1) are sharply kinked at the crossover point to make a U-turn. In contrast to the conventional co-axially stacked model, the pairwise stackings of arms 1 on 4 and 2 on 3 are interrupted by protein residues protruding into the crossover region of the junction (Fig. 2). The centre of the junction (branch point) is opened by endo VII into a parallelogram with cross dimensions of 14 Å between the exchanging strands and 32 Å between the non-exchanging strands (Fig. 1b). However, every base pair surrounding the branch point is intact (Fig. 1). All four arms are in standard B conformation and extend from the edges of the parallelogram nearly in a plane (Fig. 1b, c). Arms 1 and 3 (cleavage arms) are almost parallel with their helical axes ~15 Å apart. Arms 2 and 4 (non-cleavage arms) are tilted by ~25° towards the major groove (Fig. 1c). The roughly 'H'-shaped junction deviates from the square-planar conformation in the angles between adjacent arms, which are ~40° and ~135° instead of 90°. The Holliday junction structure observed in this crystal is in perfect agreement with the gel electrophoretic mobility of a four-way junction complexed with endo VII (ref. 12).

Endo VII approaches the Holliday junction from the minor-groove side of the branch point. Two long antiparallel helices (H2 from two subunits) abut the backbones of the non-exchanging strands with the negative end of the helical dipole, and stabilize the 30-Å separation between them (Fig. 2a). Residues Gln 67, His 70, Lys 71 and Arg 74 on these helices penetrate into the junction, keeping the exchanging strands 14 Å apart (Fig. 2b, c). In addition, the

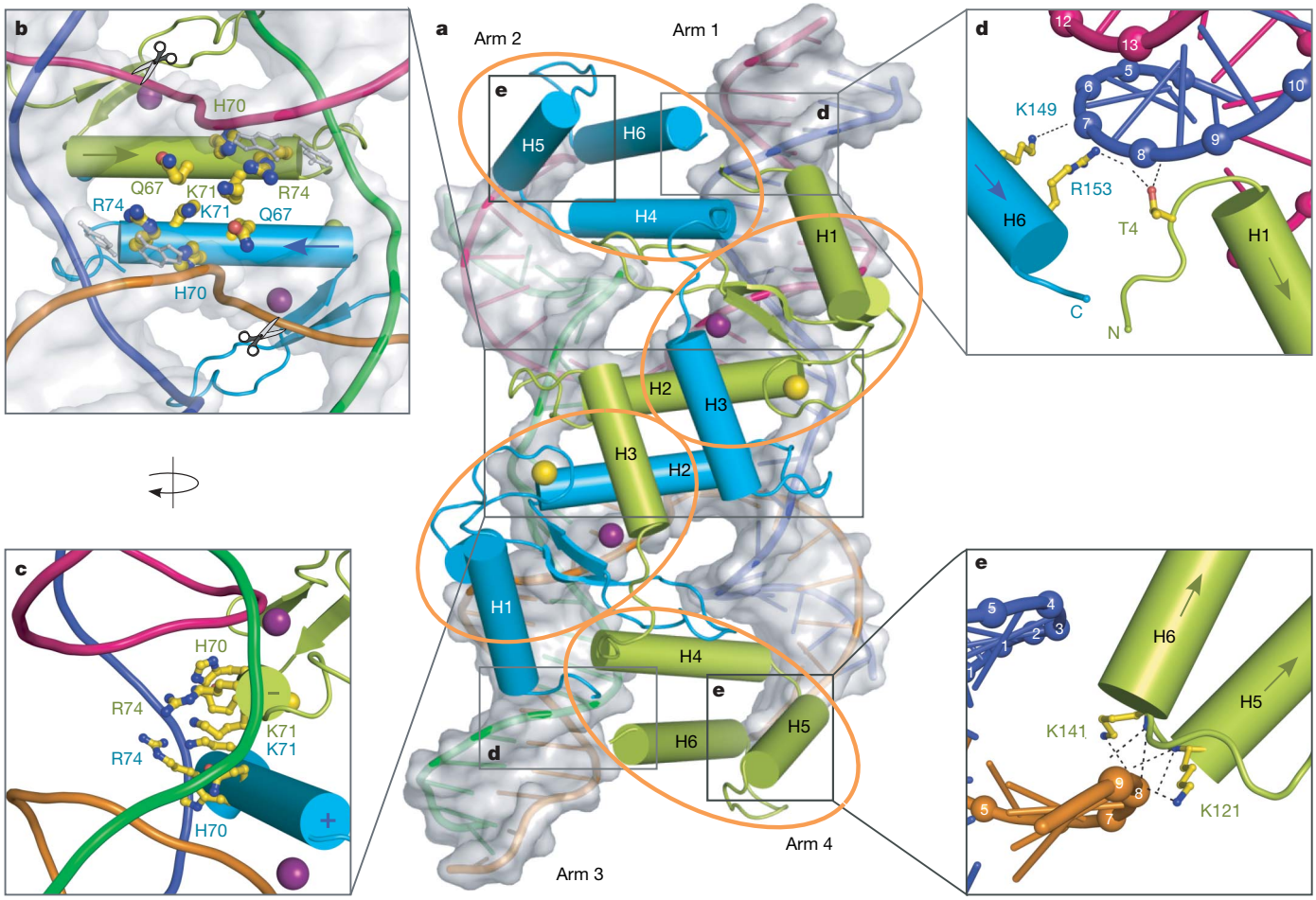


Figure 2 | Protein–DNA interactions. **a**, The back view of the complex shows endo VII in ribbon diagrams and the Holliday junction as tube-and-sticks with a semi-transparent molecular surface. The four quarters of endo VII are indicated by orange ovals, and Mg^{2+} and Zn^{2+} are shown as purple and yellow spheres, respectively. **b**, **c**, Protein–DNA interactions at the branch point. The DNA strands are colour-coded as in Fig. 1. The direction and dipolar potential of α-helices are depicted by arrows and + and – signs. **d**, **e**, Magnified views of the interactions with arm 1 (or 3) and 4 (or 2), respectively. Dashed lines represent hydrogen bonds. Base pair numbering from the junction is labelled on backbone phosphates.

©2007 Nature Publishing Group

guanidinium group of each Arg 74 stacks with the first base pair of arm 2 or 4. These are the only direct interactions between endo VII and DNA bases. Beyond the branch point, the cleavage arms are contacted by endo VII around the second and seventh base pair (Figs 2, 3a, and Supplementary Fig. 2), and arms 2 and 4 are contacted near the tenth base pair (Fig. 2e and Supplementary Fig. 2). These interactions are solely with the DNA backbones, which explains the lack of sequence specificity in endo VII cleavage.

The junction has an overall two-fold symmetry (Fig. 1), with its centre coinciding with the pseudo-dyad axis of the endo VII dimer. The two halves of the protein–Holliday junction complex are not identical and cannot be superimposed (Supplementary Fig. 3a–c). In contrast, the structure of endo VII bound to two 16-bp heteroduplexes, each of which occupies the binding site of a cleavage arm (Supplementary Fig. 1), is nearly symmetric (Supplementary Fig. 3d). Structural comparison of the two endo VII–DNA complexes reveals that the S-shaped dimeric protein is divided into four structurally constant parts: two tips and two internal turning points (Supplementary Fig. 3). Each part is mainly associated with one arm of the Holliday junction (Fig. 2a). The tip of the S, composed of three C-terminal helices (H4, H5 and H6), interacts with arms 2 and 4; the turning point of the S, composed of the N-terminal regions of one subunit and the middle region of another, interacts with arms 1 and 3 (Fig. 2). Each protein part exhibits a degree of autonomy, and the movements of the tips of the S and arms 2 and 4 are most prominent in the Holliday junction complex (Supplementary Fig. 3). Such mobility may explain the independent cleavage by two protein subunits[13]. Simultaneous interactions by helix H6 with the adjacent arms (1 and 2 or 3 and 4) 7–9 bp from the branch point may stabilize protein–DNA complexes and enable endo VII to recognize and cleave a heteroduplex as well as 3-way and 4-way junctions.

The active sites of endo VII are clearly marked by the catalytically essential residues (Asp 40, His 41 and Asn 62; refs 14, 15), the characteristic ββα fold[16] and a Mg^{2+} ion. The scissile phosphates that interact with the Mg^{2+} ions in the active sites are symmetrically located on each exchanging strand two nucleotides 3′ to the branch point. The Mg^{2+} ion is coordinated by the side chains of D40 and N62(D), the O3′ (leaving group) and the non-bridging Sp oxygen of the scissile phosphate (Fig. 3a). The location and coordination geometry of the Mg^{2+} resemble the B metal ion in a two-metal ion mechanism[17]. His 41 and His 43 are also required for efficient catalysis[14]. His 41 approaches the 5′ side of the scissile phosphate, and His 43 seems to be on the O3′ leaving-group side potentially involved in Mg^{2+} coordination. A similar arrangement of active site residues has been found in homing endonucleases I-HmuI and I-PpoI (ref. 18; Fig. 3b, c), which, like endo VII, belong to the ββα-metal family (also known as HNH endonucleases)[16]. A general mechanism for catalysis by this group of enzymes might involve His 41 (or an equivalent residue) as a general base to activate water for nucleophilic attack, whereas the divalent cation probably destabilizes the substrate, stabilizes the phosphoanion transition state and facilitates product formation.

Holliday junctions were previously crystallized either alone in the stacked-X form or in the square-planar conformation when complexed with the tetrameric RuvA or site-specific tyrosine recombinases Cre, Flp and phage λ integrase[19–24] (Fig. 4a–d). The junctions in co-crystal structures are open at the centre, as is observed with endo VII. RuvA stabilizes the Holliday junction in a perfect square

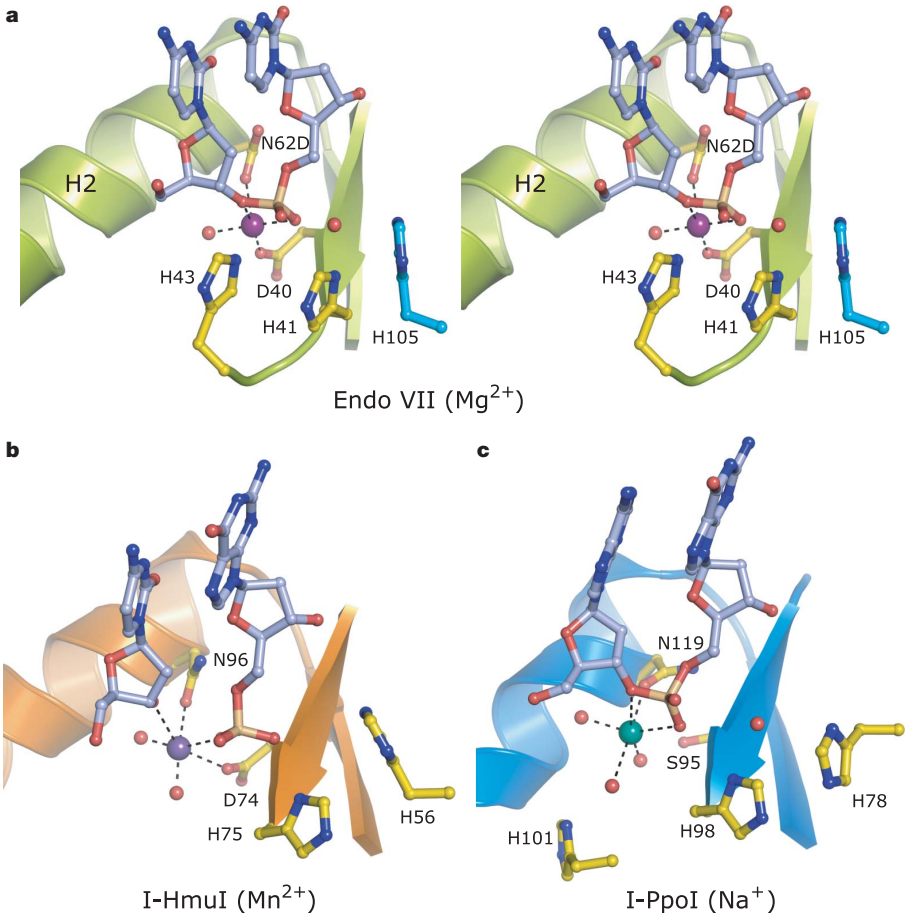


Figure 3 | The active site of endo VII. **a**, A stereo view of the active site. The Mg^{2+} ion (purple sphere) is coordinated by D40, N62D, two oxygen atoms of the scissile phosphate and a water molecule (red sphere). H41 and possibly H105 from the neighbouring subunit activate the water molecule for nucleophilic attack. **b**, The active site of I-HmuI complexed with Mn^{2+} and the DNA cleavage product (1U3E). **c**, The active site of I-PpoI complexed with DNA substrate in the presence of Na^{+} (1CZ0). The structures are identified by their Protein Data Bank (PDB) codes.

©2007 Nature Publishing Group

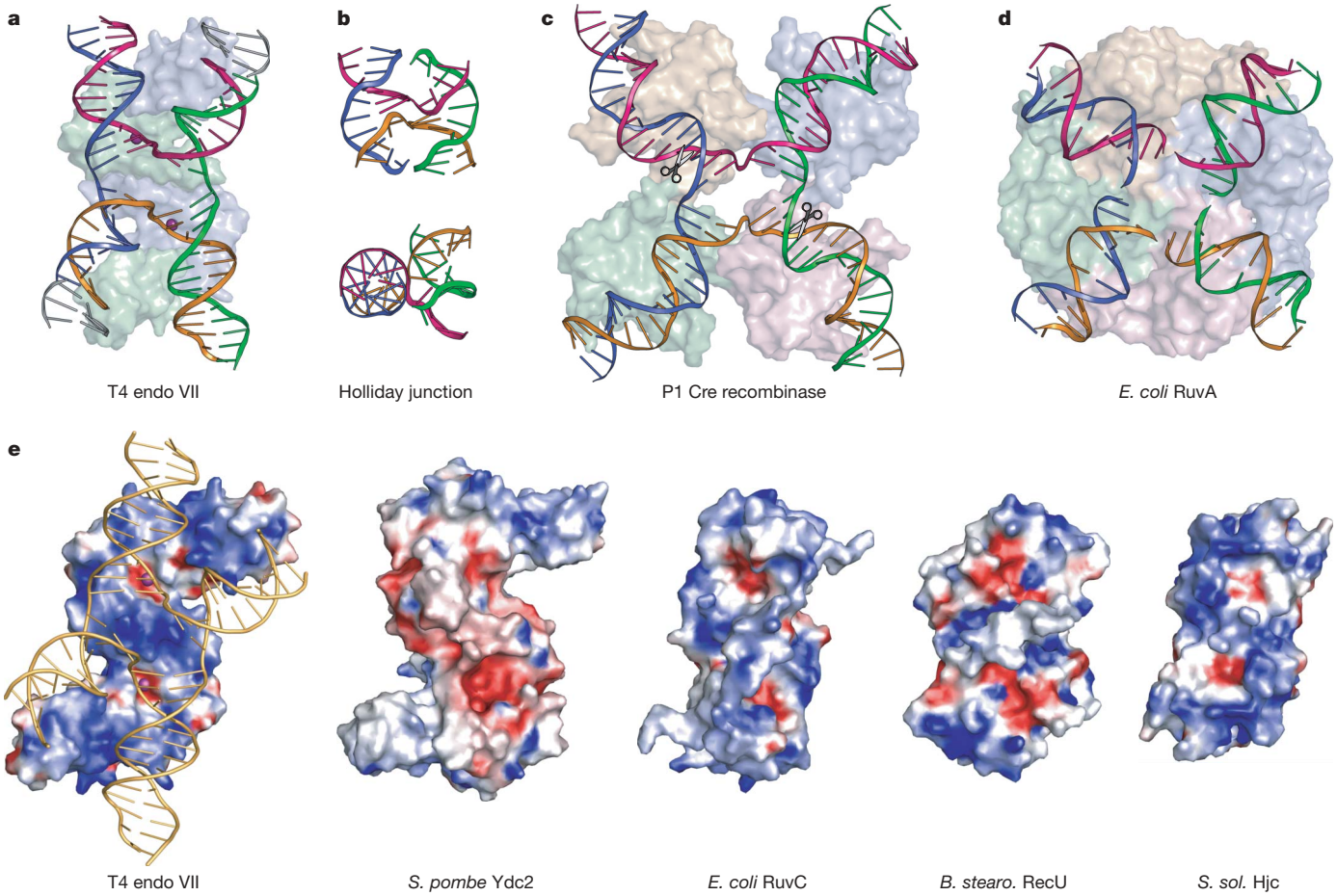


Figure 4 | Structural comparisons of Holliday junction and Holliday junction resolvases. **a**, The Holliday junction complexed with T4 endo VII. **b**, Orthogonal views of a Holliday junction in a stacked-X form (1M6G). **c**, The Holliday junction complexed with Cre recombinase in a square-planar conformation (3CRX). Each subunit is shown in a different colour. The two catalytic subunits are indicated by scissors. The N-terminal domains of Cre are removed for clarity. **d**, Four DNA duplexes complexed with *E. coli* RuvA mimic a square-planar Holliday junction (1C7Y). **e**, Molecular shape and electrostatic surface potential (blue, positive; red, negative) of Holliday junction resolvases: T4 endo VII, *S. pombe* Ydc2 (1KCF), *E. coli* RuvC (1HJR), *Bacillus stearothermophilus* (*B. stearo.*) RecU (2FCO) and *Sulfolobus solfataricus* (*S. sol.*) Hjc (1HH1).

plane, presumably to facilitate branch migration. However, in the complexes with tyrosine recombinases, the Holliday junction has little interaction with the proteins at the branch point and exhibits two-fold instead of four-fold symmetry. The opening in the centre is due to sequence-specific interactions of the arms with the four protein subunits, which pull the junction apart as dictated by the inter-subunit contacts (Fig. 4c). The absence of interactions with protein at the branch point allows the junction to isomerize during the recombination process, so each pair of strands can be sequentially acted on[23] (Supplementary Fig. 4).

Holliday junctions have now also been solved in complex with T7 endonuclease I (ref. 25), which has a type IIP restriction-endonuclease-like active site[26] and may represent a class of nucleases that recognize Holliday junctions in a stacked-X form (Fig. 4b). However, many Holliday junction resolvases, including *Schizosaccharomyces pombe* Ydc2, *Escherichia coli* RuvC, archaeal Hjc and bacterial RecU, probably share a similar substrate recognition mode with endo VII. Some of them show sequence preferences, which may be attributed to an indirect-readout mechanism[19]. Despite the lack of conservation in tertiary and quaternary structures, diverse active site compositions and different catalytic mechanisms[4], these Holliday junction resolvases have an S shape or an S-shaped positive electrostatic surface potential with comparable dimensions to those of endo VII (refs 27–30; Fig. 4d). They all possess protruding residues in the centre of the 'S', and the active sites are marked by the negative charges for metal-ion coordination near the turning points of the 'S'. These similarities suggest that they may recognize and resolve Holliday junctions in a manner similar to endo VII.

METHODS SUMMARY

T4 endo VII was expressed and purified as reported previously[8]. DNAs were prepared by annealing purified synthetic oligos. Protein and DNA samples were mixed at a ratio of one dimer per Holliday junction at a final protein concentration of 10 mg ml^{-1}. Crystals were grown by the vapour-diffusion method in 0.1 M Bis-Tris (pH 5.5–6.5), 25% (w/v) PEG 3350 and 0.1–0.2 M salt. Diffraction data were collected at ESRF and APS. The structures were solved by molecular replacement. Details of the whole process of sample preparation, diffraction data collection, structure determination and analysis are described in the Methods section.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 June; accepted 7 August 2007.
Published online 16 September 2007.

1. Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **5**, 282–304 (1964).
2. Liu, Y. & West, S. C. Happy Hollidays: 40th anniversary of the Holliday junction. *Nature Rev. Mol. Cell Biol.* **5**, 937–944 (2004).
3. Mizuuchi, K., Kemper, B., Hays, J. & Weisberg, R. A. T4 endonuclease VII cleaves Holliday structures. *Cell* **29**, 357–365 (1982).
4. Lilley, D. M. & White, M. F. The junction-resolving enzymes. *Nature Rev. Mol. Cell Biol.* **2**, 433–443 (2001).
5. Kemper, B. & Janz, E. Function of gene 49 of bacteriophage T4. I. Isolation and biochemical characterization of very fast-sedimenting DNA. *J. Virol.* **18**, 992–999 (1976).

©2007 Nature Publishing Group

6. Kemper, B. & Brown, D. T. Function of gene 49 of bacteriophage T4. II. Analysis of intracellular development and the structure of very fast-sedimenting DNA. *J. Virol.* **18,** 1000–1015 (1976).
7. Solaro, P. C., Birkenkamp, K., Pfeiffer, P. & Kemper, B. Endonuclease VII of phage T4 triggers mismatch correction *in vitro*. *J. Mol. Biol.* **230,** 868–877 (1993).
8. Raaijmakers, H. *et al.* X-ray structure of T4 endonuclease VII: a DNA junction resolvase with a novel fold and unusual domain-swapped dimer architecture. *EMBO J.* **18,** 1447–1458 (1999).
9. Raaijmakers, H., Törö, I., Birkenbihl, R., Kemper, B. & Suck, D. Conformational flexibility in T4 endonuclease VII revealed by crystallography: implications for substrate binding and cleavage. *J. Mol. Biol.* **308,** 311–323 (2001).
10. Duckett, D. R. *et al.* The structure of the Holliday junction, and its resolution. *Cell* **55,** 79–89 (1988).
11. Lilley, D. M. Structures of helical junctions in nucleic acids. *Q. Rev. Biophys.* **33,** 109–159 (2000).
12. Pöhler, J. R., Giraud-Panis, M. J. & Lilley, D. M. T4 endonuclease VII selects and alters the structure of the four-way DNA junction; binding of a resolution-defective mutant enzyme. *J. Mol. Biol.* **260,** 678–696 (1996).
13. Giraud-Panis, M. J. & Lilley, D. M. Near-simultaneous DNA cleavage by the subunits of the junction-resolving enzyme T4 endonuclease VII. *EMBO J.* **16,** 2528–2534 (1997).
14. Giraud-Panis, M. J. & Lilley, D. M. T4 endonuclease VII. Importance of a histidine-aspartate cluster within the zinc-binding domain. *J. Biol. Chem.* **271,** 33148–33155 (1996).
15. Golz, S., Christoph, A., Birkenkamp-Demtröder, K. & Kemper, B. Identification of amino acids of endonuclease VII essential for binding and cleavage of cruciform DNA. *Eur. J. Biochem.* **245,** 573–580 (1997).
16. Scholz, S. R. *et al.* Experimental evidence for a ββα-Me-finger nuclease motif to represent the active site of the caspase-activated DNase. *Biochemistry* **42,** 9288–9294 (2003).
17. Yang, W., Lee, J. Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg^{2+}-ion catalysis and substrate specificity. *Mol. Cell* **22,** 5–13 (2006).
18. Stoddard, B. L. Homing endonuclease structure and function. *Q. Rev. Biophys.* **38,** 49–95 (2005).
19. Khuu, P. A., Voth, A. R., Hays, F. A. & Ho, P. S. The stacked-X DNA Holliday junction and protein recognition. *J. Mol. Recognit.* **19,** 234–242 (2006).
20. Hargreaves, D. *et al.* Crystal structure of *E. coli* RuvA with bound DNA Holliday junction at 6 Å resolution. *Nature Struct. Biol.* **5,** 441–446 (1998).
21. Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H. & Morikawa, K. Crystal structure of the Holliday junction DNA in complex with a single RuvA tetramer. *Proc. Natl Acad. Sci. USA* **97,** 8257–8262 (2000).
22. Gopaul, D. N., Guo, F. & Van Duyne, G. D. Structure of the Holliday junction intermediate in Cre-*loxP* site-specific recombination. *EMBO J.* **17,** 4175–4187 (1998).
23. Conway, A. B., Chen, Y. & Rice, P. A. Structural plasticity of the Flp–Holliday junction complex. *J. Mol. Biol.* **326,** 425–434 (2003).
24. Biswas, T. *et al.* A structural basis for allosteric control of DNA recombination by λ integrase. *Nature* **435,** 1059–1066 (2005).
25. Hadden, J. M., Déclais, A.-C., Carr, S. B., Lilley, D. M. J. & Phillips, E. V. The structural basis of Holliday junction resolution by T7 endonuclease I. *Nature* doi:10.1038/nature06158 (this issue).
26. Déclais, A. C., Hadden, J., Phillips, S. E. & Lilley, D. M. The active site of the junction-resolving enzyme T7 endonuclease I. *J. Mol. Biol.* **307,** 1145–1158 (2001).
27. Ariyoshi, M. *et al.* Atomic structure of the RuvC resolvase: a Holliday junction-specific endonuclease from *E. coli*. *Cell* **78,** 1063–1072 (1994).
28. Bond, C. S., Kvaratskhelia, M., Richard, D., White, M. F. & Hunter, W. N. Structure of Hjc, a Holliday junction resolvase, from *Sulfolobus solfataricus*. *Proc. Natl Acad. Sci. USA* **98,** 5509–5514 (2001).
29. Ceschini, S. *et al.* Crystal structure of the fission yeast mitochondrial Holliday junction resolvase Ydc2. *EMBO J.* **20,** 6601–6611 (2001).
30. McGregor, N. *et al.* The structure of *Bacillus subtilis* RecU Holliday junction resolvase and its role in substrate selection and sequence-specific cleavage. *Structure* **13,** 1341–1351 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Leahy and M. Gellert for critical reading of the manuscript, S. Ramon-Maiques for collecting the diffracting data of the endo VII–heteroduplex complex, and D. M. Lilley for background reading materials. C.B. thanks J. Basquin, E. Ennifar and C. Sauter for help with crystallization and data collection, and M. Nowotny and J. Y. Lee for help with manuscript preparation. This research was supported by EMBL, Deutsche Forschungsgemeinschaft and the Intramural Research Program of NIDDK, NIH.

Author Contributions C.B. carried out all experiments. The project was initiated at EMBL and finished at NIH. All authors contributed to experimental design, interpretation and manuscript preparation.

Author Information Atomic coordinates and structure factors of the endo VII–DNA complexes have been deposited in the Protein Data Bank. The accession codes are 2QNC and 2QNF for the Holliday junction and the heteroduplex complex, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to W.Y. (wei.yang@nih.gov) or D.S. (suck@embl.de).

©2007 Nature Publishing Group

METHODS

Protein and DNA sample preparation. T4 endo VII was expressed and purified as reported previously[8] with minor modifications. The purified protein was stored in 10 mM HEPES (pH 7.0), 250 mM NaCl and 1 mM DTT with or without 10 mM $MgCl_2$. Oligonucleotides were either purchased in HPLC-purified form from Metabion or were obtained from the Facility for Biotechnology Resources (FDA/NIH, Bethesda) and subsequently HPLC purified on a Poros oligo R3 column. Sequences of DNA oligos in the co-crystals are diagrammed in Supplementary Fig. 2. To make the Holliday junction of 4w214p (2 bp and 14 bp in alternating arms), two pairs of oligos—4w214p-1 (CACATCGATGGAGCCG) and 4w214p-2 (CGGCTCCATCGATGTG), and 4w214p-3 (CACATCGATGGAGCGC) and 4w214p-4 (GCGCTCCATCGATGTG)—were annealed separately before mixing at a 1:1 molar ratio. To make the Holliday junction of 4w1014 (10 and 14 bp in alternating arms), 4w1014-1 (CGAAGAATTCCGGATTAGGGATCC), 4w1014-2 (GGATCCCTAAGCTCCATCGATGTG), 4w1014-3 (CACATCGATGGAGCCGCTAGGCCT) and 4w1014-4 (AGGCCTAGCGTCCGGAATTCTTCG) were mixed at an equimolar ratio and were annealed by heating to 363 K for 5 min and slow-cooling to 295 K in 90 min. Protein and DNA samples were mixed at a ratio of one dimer per junction at a final protein concentration of 10 mg ml^{-1}.

Crystallization. Crystals were grown by the vapour-diffusion method at 293 K. Initial crystallization conditions were found using the Index Screen (Hampton Research). Crystals of endo VII–4w214p complexes were obtained in 25% (w/v) PEG 3350, 100 mM Bis-Tris (pH 6.5), 220 mM Li_2SO_4 and 5% (w/v) glucose and were transferred to a stabilization solution containing 20% (w/v) PEG 3350, 100 mM Bis-Tris/HCl (pH 6.5), 180 mM Li_2SO_4 and 20%(w/v) glucose. The crystals of endo VII–4w1014 complexes were grown in 0.1 M Bis-Tris (pH 5.5) and 25% (w/v) PEG 3350, and were stepwise soaked in a reservoir solution containing 5, 10 and 20% (w/v) ethylene glycol for cryoprotection.

Diffraction data collection and structure determination. Diffraction data of the endo VII–Holliday junction crystals were collected at ID 14-1 (ESRF) after they were flash-cooled in liquid nitrogen and processed with XDS[31]. Data could be processed equally well in space groups $P4_3$ and $P4_32_12$. Twinning analysis indicated partial merohedral twinning for the lower symmetry space group. Diffraction data of the endo VII–4w214p crystals were collected at 22-ID (APS) at 100 K and processed with HKL2000 (ref. 32). The crystal belonged to space group $P3_1$. Structure factors for both data sets were calculated with TRUNCATE[33] (Supplementary Table 1).

Molecular replacement with PHASER[34] was successfully performed on the endo VII–4w214p crystal using an apo protein structure (PDB code 1E7D) as a search model. Because crystals diffracted anisotropically, the anisotropic correction function in PHASER was used. The resulting electron density map was of a poor quality and was improved after density modification by solvent flipping using CNS[35]. Two DNA duplexes were clearly visible in the $2F_o - F_c$ map. The model of protein and DNA was built and refined up to 3 Å by multiple rounds of manual fitting in O[36] or COOT[37] and positional refinement in CNS. Interestingly, the DNA did not form a Holliday junction of alternating 2-bp and 14-bp arm lengths inside the crystal. Instead, two 16-mer duplexes with two mismatching base pairs each are associated with an endo VII dimer in the asymmetric unit (Supplementary Fig. 1).

The protein part of the refined endo VII–4w214p complex was subsequently used to solve the structure of the Holliday junction complex by PHASER, which found a solution only in the space group $P4_3$. After applying density modification (solvent flipping) in CNS, one arm of the Holliday junction was clearly visible in a $2F_o - F_c$ map. After cycles of model building and refinement, the entire Holliday junction was built in. The final refinement was carried out using data processed in space group $P4_32_12$, of which each asymmetric unit contains one endo VII dimer and one Holliday junction. Because of the rather low resolution, it was not clear which orientation the DNA has inside the crystal. Refinement was also carried out with the DNA sequence switched between arms 1 and 3, and 2 and 4, which resulted in similar refinement statistics. We assume that there is a mixture of both orientations in the crystals. Because there is no sequence-specific recognition, we used one orientation for refinement to keep the parameter-to-data ratio reasonable.

DNA structures and protein–DNA interactions were analysed using the programs CURVES[38] and NUCPLOT[39]. Structural superimpositions were performed using the program LSQMAN[40]. Electrostatic surface potentials were calculated using GRASP[41], and all molecular visualizations were prepared using PYMOL (http://pymol.sourceforge.net).

31. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26,** 795–800 (1993).
32. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
33. The. CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
34. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).
35. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
36. Jones, T. A., Zou, J.-Y. & Cowan, S. W. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
37. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
38. Lavery, R. & Sklenar, H. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* **6,** 655–667 (1989).
39. Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. NUCPLOT: a program to generate schematic diagrams of protein–nucleic acid interactions. *Nucleic Acids Res.* **25,** 4940–4945 (1997).
40. Kleywegt, G. J. Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D* **52,** 842–857 (1996).
41. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11,** 281–296 (1991).

©2007 Nature Publishing Group

LETTERS

The structural basis of Holliday junction resolution by T7 endonuclease I

Jonathan M. Hadden[1], Anne-Cécile Déclais[2], Stephen B. Carr[1], David M. J. Lilley[2] & Simon E. V. Phillips[1]

The four-way (Holliday) DNA junction is the central intermediate in homologous recombination, a ubiquitous process that is important in DNA repair and generation of genetic diversity[1]. The penultimate stage of recombination requires resolution of the DNA junction into nicked-duplex species by the action of a junction-resolving enzyme, examples of which have been identified in a wide variety of organisms[2]. These enzymes are nucleases that are highly selective for the structure of branched DNA. The mechanism of this selectivity has, however, been unclear in the absence of structural data. Here we present the crystal structure of the junction-resolving enzyme phage T7 endonuclease I in complex with a synthetic four-way DNA junction. Although the enzyme is structure-selective, significant induced fit occurs in the interaction, with changes in the structure of both the protein and the junction. The dimeric enzyme presents two binding channels that contact the backbones of the junction's helical arms over seven nucleotides. These interactions effectively measure the relative orientations and positions of the arms of the junction, thereby ensuring that binding is selective for branched DNA that can achieve this geometry.

Phage T7 endonuclease I (endo I) is a basic protein of 149 amino acids[3–5]. It is a member of the nuclease superfamily[6], with an active site similar to that of a number of restriction enzymes, particularly *BglI* (ref. 7). Endo I binds to four-way DNA junctions as a dimer with a dissociation constant of ~1 nM (ref. 8) and cleaves the two continuous DNA strands adjacent to the crossover point. The complexes are not displaced by 1,000-fold higher concentrations of duplex DNA[9].

Previously, we have reported the crystal structures of a catalytically impaired mutant of endo I(E65K)[7] and of the wild-type enzyme in the presence of Mn^{2+} ions[10], both in the absence of DNA. Endo I forms an intimately associated symmetrical homodimer comprising two domains, each formed by residues 17–44 from one subunit and 50–145 from the other. The domains are connected by a bridge that forms part of an extended β-sheet. On the basis of these structures, together with that of *BglI* bound to DNA[11] and biochemical data, we constructed a model of the complex[12]. In this, pairs of DNA helical arms are approximately aligned end-to-end along the same axis (coaxially aligned) and the two pairs of arms lie approximately perpendicular to each other. Two alternative forms, or conformers, of the complex exist in which different pairs of arms coaxially align. We have shown that by replacing two of the DNA helical arms of the Holliday junction with short hairpin loops it is possible to bias the complex strongly to one conformer[8]. We reasoned that this bias might aid crystallization. Using a junction comprising diagonally opposed stem-loops of 5 base pairs (bp) (Fig. 1), we obtained crystals of the complex with a catalytically impaired mutant of the protein (K67A) in the presence of calcium (Ca^{2+}) ions to prevent cleavage during crystallization. We have solved the structure of this complex using molecular replacement, and it has been refined to 3.1 Å resolution with good statistics (R_{factor}, 26.7%; R_{free}, 29.8%).

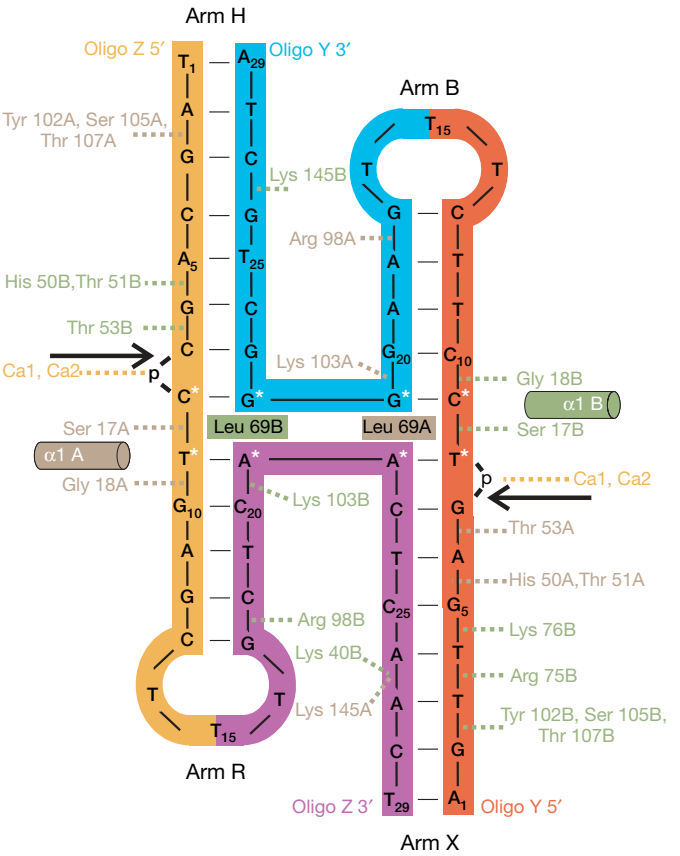


Figure 1 | Schematic representation of the Holliday junction in the crystal structure. The junction comprises four arms labelled B, H, R and X, as shown; arms B and R are 5-bp stem-loops. The constituent oligonucleotides are named Y (with its 5′ end in arm X) and Z (5′ end in arm H) and the sequences are coloured to match Figs 2 and 3. The continuous strands of the DNA junction are depicted in gold and red whereas the exchanging strands are in purple and cyan. The hairpin loops at the ends of arms B and R are not visible in electron density maps. Amino acids interacting with the DNA have been coloured according to their protein chain (protein chain A in fawn and B in pale green) (see also Supplementary Table 2). Each active site of the protein binds two calcium ions (Ca1 and Ca2), which also interact with the DNA backbone. Dipoles of α1-helices (residues 17–29; cylinders) of both protein subunits make additional interactions with the DNA. Scissile phosphates are shown (p), with scissile bonds indicated by arrows, and loss of base-stacking near the core of the junction by asterisks.

[1]Astbury Centre for Structural Molecular Biology, Institute of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. [2]Cancer Research UK Nucleic Acid Structure Research Group, MSI/WTB complex, University of Dundee, Dundee DD1 4HN, UK.

©2007 Nature Publishing Group

The overall structure of the complex (Fig. 2) is in excellent agreement with the earlier model. The pairs of DNA helical arms are essentially coaxial (despite the central bases being unstacked, see below), with an interaxial angle of –80°(defined as the minimum rotation required to superimpose the axis of duplex HR on that of BX, viewed from HR to BX, with clockwise rotation positive). Thus, the angle between the axes has been altered by ~130°, from the right-handed stacked X-structure of the free Holliday junction in solution (interaxial angle is approximately +50°, refs 13, 14) into an approximately perpendicular structure. This is consistent with studies of the structure of the endo I–Holliday junction complex in solution[12]. The observed conformation creates an almost continuous deep groove across the major groove face of the junction (Fig. 2a and Supplementary Movie 1), which forms the principal binding site for the protein. As a consequence, the minor groove face of the junction is splayed out and is virtually protein-free in this structure, although amino acids 1–15 of the protein cannot be located in the electron density map. The structure of the endo I complex contrasts with those of other junction–enzyme complexes, which in general adopt approximately planar structures with loss of coaxial alignment of arms[15–23]. This is well illustrated by the only other junction-resolving enzyme complex structure, T4 endonuclease VII (ref. 16).

The conformation of the DNA in the complex brings the phosphate backbone of the continuous strands into close proximity with the two active sites of the dimeric enzyme. This structure provides the first picture of the DNA substrate bound into the active site of endo I (Fig. 2d). Two Ca^{2+} ions are bound in each active site, and these are coordinated by the carboxylate side chains of Asp 55 and Glu 65. The active site corresponds well to the P-D…D/E-X-K consensus sequence of a number of type 2 restriction enzymes. The Ca^{2+} ions have been refined at occupancies of 100% and 75%, as previously described in our DNA-free structure[10]; these occupancy levels are consistent with our current electron density maps. We expect that the scissile phosphate is hydrolysed on nucleophilic attack by a water molecule that is directly coordinated to the full occupancy metal ion (which would be Mg^{2+} in the active enzyme) bound at site one[24].

Although the nuclease function is clearly related to that of restriction enzymes, the key distinguishing feature of the junction-resolving enzymes is their strong selectivity for the structure of the DNA branchpoint. Relative to the free protein, the two domains of endo I exhibit a re-orientation relative to one another on binding the junction, corresponding to a rotation of 18° and bending about the protein's two-fold axis (Supplementary Movie 2). This creates two almost perpendicular channels in the protein of ~30 Å in length, each formed by the front of one domain and the back of the other (Fig. 2b, c). These positively charged channels form extensive contacts with the arms containing the 5′ ends of the continuous strands (Fig. 1) and grip the DNA duplexes in a specific relative orientation, in which the angle between these two duplexes is ~80°. This is the basis for the selectivity for a junction—only DNA structures that can adopt this shape will be tightly bound[25]. Binding the arms of the junction into these channels results in the burial of 4,180 $Å^2$ solvent-accessible protein surface. This intimate association is consistent with the extremely slow dissociation rate observed for the complex[8]; this might help to prevent premature dissociation before both cleavages of the DNA strands have been made, consistent with earlier data[26]. The interaction is mediated by numerous hydrogen bonds (Fig. 1),

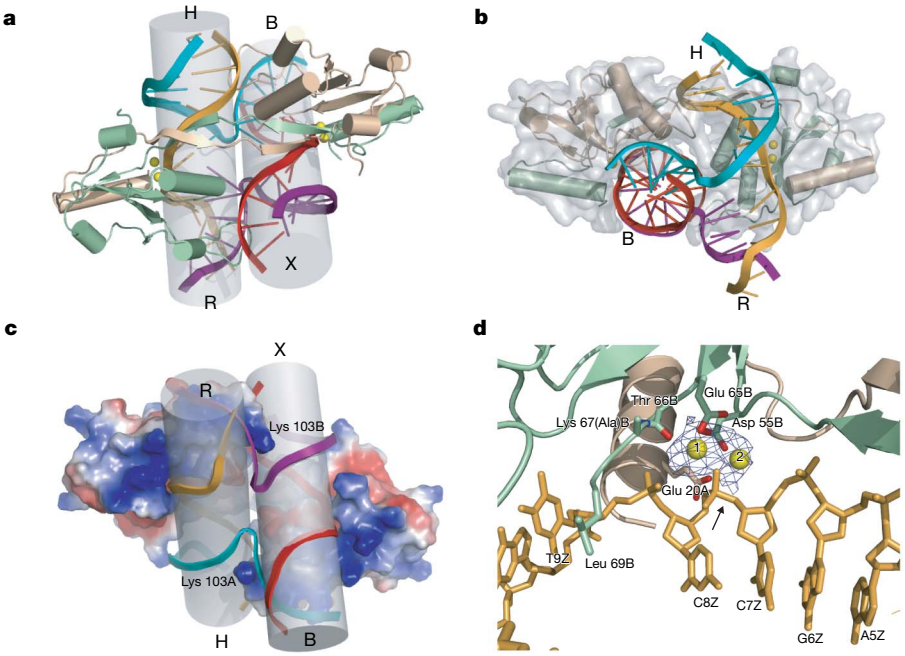


Figure 2 | The structure of the complex and the active site. The two polypeptides are shown in pale green and fawn. The Ca^{2+} ions bound at the active sites are shown in yellow. **a**, View of the major groove side of the junction. The pairwise coaxial alignment of DNA helices is highlighted by the grey cylinders. The protein is bound in the continuous widened major groove, and the active sites (indicated by the Ca^{2+} ions) are located on the continuous strands. **b**, The DNA-binding channels of the protein, viewed along the axis of one pair (X and B) of DNA helices. The near-perpendicular protein channels are 30 Å in length and wrap around the DNA helices. **c**, Surface representation of the complex coloured according to electrostatic charge distribution, showing the opposite face from that in **a**. The scale is +80 (blue) to –80 (red) $kT\,e^{-1}$. Lys 103 forms a positively charged finger, which reduces electrostatic repulsion between phosphate groups forming the tight turn at the point of strand exchange. **d**, The active site containing arm H of the DNA junction (oligo Z nucleotides 5–9), with the scissile phosphate bond shown by the arrow. The metal ion at site 1 is coordinated by Asp 55, Glu 65 and Thr 66, whereas that at site 2 is coordinated by Asp 55. The electron density shown is an $F_o - F_c$ map calculated in the absence of metal ions. An additional ligand to both metals is provided by oxygen O1 of the 5′ phosphate group of nucleotide C8Z (see coordination distances in Supplementary Table 3).

©2007 Nature Publishing Group

often involving charged amino acids (Supplementary Table 2). The protein interacts exclusively with the backbone of the DNA and makes no hydrogen bonds with the bases. Previously, we have shown that endo I exhibits significant discrimination for the sequence of the arms[8], and the crystallographic data indicate that this arises only from an indirect read-out mechanism resulting from subtle variations in the conformation of the phosphate backbone.

Binding of endo I results in significant distortion of the structure of the junction, both globally (discussed above) and locally. The DNA backbone makes a tight turn at the point of strand exchange. The adjacent phosphate groups on the exchanging strand come within ~6 Å of one another in this region and the side chain of Lys 103 is inserted between these two groups (Figs 1, 2c and 3), helping to reduce the electrostatic repulsion at the centre of the junction. The electron density maps in the central region of the junction reveal a very open structure (Fig. 3) in which the central base pairs on the continuous strands are completely unstacked, with the base–base distance (nucleotides 8–9 in each chain) increasing to ~10 Å in the complex. This is consistent with the reactivity of bases in this region to permanganate oxidation and the enhanced fluorescence intensity of 2-aminopurine placed in this location[12].

There is also evidence for an additional dynamic character of bases in this region. Electron density can be observed in both stacked and un-stacked positions for nucleotides Guanine 21 in oligo Y and Adenine 21 in oligo Z (see Fig. 1 for numbering scheme), located at the point of strand exchange in arms B and R (Fig. 3). This indicates a dynamic equilibrium between equally populated alternative conformations, in which the nucleobases undergo transitions between a base-paired and an unpaired conformation where the purine is displaced towards the centre of the junction. The side chain of Leu 69 projects into the hydrophobic core formed by the stacked bases, and packs against bases Cytidine 9 in oligo Y and Thymine 9 in oligo Z (Fig. 3). The conformations of nucleobases at the point of strand exchange may offer some mechanistic insight into potential early steps in the process of branch migration.

It seems improbable that endo I initiates the gross structural re-orientation observed in our crystal structure. Rather, it is more likely that the protein recognizes and binds tightly to a transiently populated DNA species. This is consistent with the slow rates of association we have measured[8]. In free solution, DNA junctions are in dynamic exchange between stacking conformers[27]; analysis of the process in single junction molecules has indicated that an open structure is an intermediate in their interconversion[28]. Once unstacked, the junction would be extremely flexible and the large surface area of endo I available for interaction would serve to stabilize this intermediate species, which may then be further deformed by the protein at little energetic cost.

The junction-resolving enzymes are models for the molecular recognition of DNA structure on a large scale. The structure of the complex of endo I bound to a DNA junction shows how the geometry of this branched structure can be recognized, while at the same time being distorted by the enzyme. The recognition process exploits the dynamic character of the junction, allowing it to be 'moulded' onto the large binding surface of the protein. Although functionally related enzymes including endonuclease VII may differ in the manner of their interactions, the underlying principles of recognition are likely to be similar.

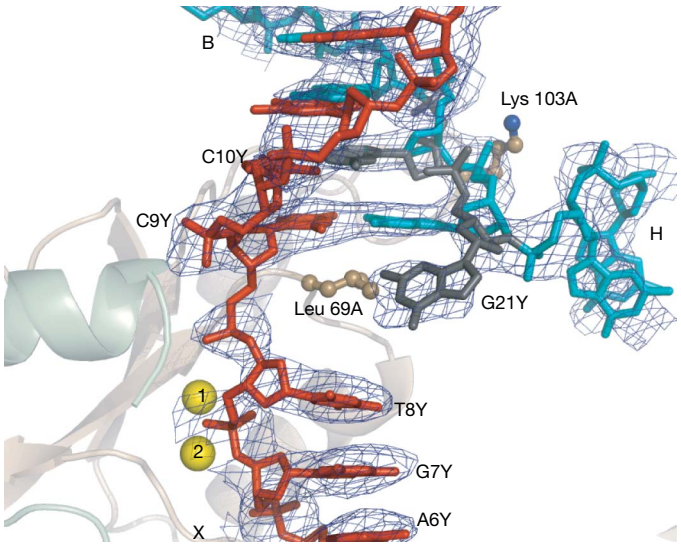


Figure 3 | Local distortion at the branch point. View of arm B of the junction showing the two conformations (cyan and grey) observed for G21 in oligo Y. This base either pairs with C9 in oligo Y or is unpaired and directed towards the centre of the complex. Leu 69 caps the end of one of the DNA helices. The $2F_o - F_c$ electron density map is displayed around the DNA. Oligo Z has been omitted for clarity. Metal ions are shown as yellow spheres.

METHODS SUMMARY

Endo I(K67A) was expressed in *Escherichia coli* and purified using Ni-NTA and anion-exchange chromatography. Two oligonucleotides (Y and Z), each containing a 3-thymine hairpin-loop, were annealed to form a four-way junction. The endo I–Holliday junction complex was formed by mixing the protein and nucleic acid in a 1:1.2 molar ratio (protein:DNA) and purifying the complex using gel filtration chromatography. Crystals of endo I–Holliday junction complex were grown using the sitting-drop vapour diffusion method with a precipitant containing 13–16% PEG 6000, 30 mM $CaCl_2$, 100 mM MES, pH 6.7. Diffraction data were collected at station ID23-2 at the European Synchrotron Radiation Facility (ESRF), and the structure of the complex was solved by molecular replacement using the structure of free endo I (PDB Entry 1M0D) as the search model. All crystallographic procedures were performed using programs from the CCP4 suite (ref. 29) and figures were prepared using PyMOL (http://www.pymol.org).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 May; accepted 7 August 2007.
Published online 16 September 2007.

1. Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **5,** 282–304 (1964).
2. Lilley, D. M. & White, M. F. The junction-resolving enzymes. *Nature Rev. Mol. Cell Biol.* **2,** 433–443 (2001).
3. Center, M. S. & Richardson, C. C. An endonuclease induced after infection of *Escherichia coli* with bacteriophage T7. I. Purification and properties of the enzyme. *J. Biol. Chem.* **245,** 6285–6291 (1970).
4. Sadowski, P. D. Bacteriophage T7 endonuclease. I. Properties of the enzyme purified from T7 phage-infected *Escherichia coli* B. *J. Biol. Chem.* **246,** 209–216 (1971).
5. de Massey, B., Studier, F. W., Dorgai, L., Appelbaum, F. & Weisberg, R. A. Enzymes and the sites of genetic recombination: studies with gene-3 endonuclease of phage T7 and with site-affinity mutants of phage. *Cold Spring Harb. Symp. Quant. Biol.* **49,** 715–726 (1984).
6. Lilley, D. M. & White, M. F. Resolving the relationships of resolving enzymes. *Proc. Natl Acad. Sci. USA* **97,** 9351–9353 (2000).
7. Hadden, J. M., Convery, M. A., Déclais, A. C., Lilley, D. M. & Phillips, S. E. Crystal structure of the Holliday junction resolving enzyme T7 endonuclease I. *Nature Struct. Biol.* **8,** 62–67 (2001).
8. Déclais, A.-C., Liu, J., Freeman, A. D. & Lilley, D. M. Structural recognition between a four-way DNA junction and a resolving enzyme. *J. Mol. Biol.* **359,** 1261–1276 (2006).
9. Duckett, D. R., Giraud-Panis, M.-E. & Lilley, D. M. Binding of the junction-resolving enzyme bacteriophage T7 endonuclease I to DNA: separation of binding and catalysis by mutation. *J. Mol. Biol.* **246,** 95–107 (1995).
10. Hadden, J. M., Déclais, A.-C., Phillips, S. E. & Lilley, D. M. Metal ions bound at the active site of the junction-resolving enzyme T7 endonuclease I. *EMBO J.* **21,** 3505–3515 (2002).
11. Newman, M. *et al.* Crystal structure of restriction endonuclease *Bgl*I bound to its interrupted DNA recognition sequence. *EMBO J.* **17,** 5466–5476 (1998).
12. Déclais, A.-C. *et al.* The complex between a four-way DNA junction and T7 endonuclease I. *EMBO J.* **22,** 1398–1409 (2003).
13. Ortiz-Lombardia, M. *et al.* Crystal structure of a DNA Holliday junction. *Nature Struct. Biol.* **6,** 913–917 (1999).

©2007 Nature Publishing Group

14. Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Ho, P. S. The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl Acad. Sci. USA* **97,** 3971–3976 (2000).
15. Pöhler, J. R., Giraud-Panis, M. J. & Lilley, D. M. T4 endonuclease VII selects and alters the structure of the four-way DNA junction; binding of a resolution-defective mutant enzyme. *J. Mol. Biol.* **260,** 678–696 (1996).
16. Biertümpfel, C., Yang, W., &. Suck, D. Crystal structure of T4 endonuclease VII resolving a Holliday junction. *Nature* doi:10.1038/nature06152 (this issue).
17. Gopaul, D. N., Guo, F. & Van Duyne, G. D. Structure of the Holliday junction intermediate in Cre–*loxP* site-specific recombination. *EMBO J.* **17,** 4175–4187 (1998).
18. Chen, Y., Narendra, U., Iype, L. E., Cox, M. M. & Rice, P. A. Crystal structure of a Flp recombinase–Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell* **6,** 885–897 (2000).
19. Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H. & Morikawa, K. Crystal structure of the Holliday junction DNA in complex with a single RuvA tetramer. *Proc. Natl Acad. Sci. USA* **97,** 8257–8262 (2000).
20. Roe, S. M. *et al.* Crystal structure of an octameric RuvA–Holliday junction complex. *Mol. Cell.* **2,** 361–372 (1998).
21. Hargreaves, D. *et al.* Crystal structure of *E. coli* RuvA with bound DNA Holliday junction at 6 Å resolution. *Nature Struct. Biol.* **5,** 441–446 (1998).
22. Bennett, R. J. & West, S. C. Structural analysis of the RuvC–Holliday junction complex reveals an unfolded junction. *J. Mol. Biol.* **252,** 213–226 (1995).
23. White, M. F. & Lilley, D. M. The resolving enzyme CCE1 of yeast opens the structure of the four-way DNA junction. *J. Mol. Biol.* **266,** 122–134 (1997).
24. Liu, J., Déclais, A.-C. & Lilley, D. M. Mechanistic aspects of the DNA junction-resolving enzyme T7 endonuclease I. *Biochemistry* **45,** 3934–3942 (2006).
25. Bhattacharyya, A. *et al.* Model for the interaction of DNA junctions and resolving enzymes. *J. Mol. Biol.* **221,** 1191–1207 (1991).
26. Parkinson, M. J. & Lilley, D. M. The junction-resolving enzyme T7 endonuclease I: quaternary structure and interaction with DNA. *J. Mol. Biol.* **270,** 169–178 (1997).
27. McKinney, S. A., Déclais, A.-C., Lilley, D. M. & Ha, T. Structural dynamics of individual Holliday junctions. *Nature Struct. Biol.* **10,** 93–97 (2003).
28. Joo, C., McKinney, S. A., Lilley, D. M. & Ha, T. Exploring rare conformational species and ionic effects in DNA Holliday junctions using single-molecule spectroscopy. *J. Mol. Biol.* **341,** 739–751 (2004).
29. Collaborative computational project number 4.. The *CCP*4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Flot at the ESRF for his assistance with the X-ray data collection and V. Sergeant for technical support. This work was supported by funds from the Wellcome Trust, BBSRC and Cancer Research UK.

Author Contributions J.M.H. purified and crystallized the complex, collected data, and built and refined the structure. A.-C.D. designed DNA oligonucleotides for co-crystallization and performed the molecular biology and biochemical studies. S.B.C. collected data and refined the structure of the complex. D.M.J.L. and S.E.V.P. conceived and designed the study and also helped to analyse the data. All authors discussed the results and wrote the paper.

Author Information Atomic coordinates and structure factors have been deposited in the PDB database under the accession code 2PFJ. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.E.V.P. (s.e.v.phillips@leeds.ac.uk).

©2007 Nature Publishing Group

METHODS

Protein expression and purification. Full-length T7 endo I with an amino-terminal deca-histidine affinity tag linked by a peptide containing a tobacco etch virus (TEV) protease site was expressed in *E. coli* using the expression vector pET-TEV-endoI (ref. 30). To inhibit junction cleavage but preserve the junction- and metal-binding activity of the protein, allowing preparation of enzyme–junction complexes, a mutation (K67A) was made that significantly reduces DNA cleavage activity. Endo I(K67A) was expressed and purified as previously described[30]; however, after removal of the N-terminal deca-histidine affinity tag with TEV protease, the protein was further purified using anion-exchange chromatography.

Oligonucleotide sequences. 5′-AGTTGAGTCCTTCTTTGAAGGGGCTGCTA-3′ (Y) and 5′-TAGCAGCCTGAGCTTTGCTCAACTCAACT-3′ (Z) were annealed to form a synthetic Holliday junction.

Crystallization. Before crystallization, the complex was concentrated in 50 mM Tris HCl, 50 mM NaCl, 5 mM EDTA, pH 8.0 until the protein concentration reached 5.8 mg ml^{-1}. Sitting drop vapour diffusion experiments were set up by mixing 1 μl complex solution with 1 μl precipitant solution, 13–16% PEG 6000, 30 mM $CaCl_2$, 100 mM MES, pH 6.7. Drops were equilibrated against 500 μl precipitant solution.

Data collection, processing, model building and refinement. Before collecting data, crystals of endo I–Holliday junction complex were soaked in cryoprotectant (20% PEG 6000, 30 mM $CaCl_2$, 25% ethylene glycol, 100 mM MES, pH 6.7) for 3 min. Crystals were then immediately frozen at 100 K before data collection. Data were collected at a wavelength of 0.87 Å on station ID23-2 at the ESRF. Crystals of endo I–Holliday junction complex diffracted to 3.1 Å and belong to the trigonal space group $P3_221$ with unit cell dimensions $a = b = 92.98$ Å, $c = 124.26$ Å. All data were integrated using MOSFLM[31] and were scaled using SCALA[32]. Data collection and processing statistics are shown in Supplementary Table 1.

The structure of the complex was solved by molecular replacement with the program PHASER[33] using the globular domain of free endo I (PDB ID: 1M0D) as the initial search model. Initial electron density maps calculated by PHASER clearly showed areas of density that could be attributed to duplex DNA. This was introduced into the model and refinement of the structure proceeded with alternating cycles of refinement using REFMAC and manual model-building using COOT. Tight non-crystallographic symmetry restraints were maintained between the protein chains throughout the refinement process and the DNA arms were tightly restrained to maintain Watson–Crick base pairing unless there was clear evidence in the electron density maps that this was not the case. The calcium–protein coordination distances were also restrained to be similar to those observed for the structure of free endo I (1M0D).

Multiple rounds of rebuilding and refinement resulted in a complete model spanning residues 16–145 for the protein and bases 1–29 in each of the oligonucleotides (with the exception of the poly-T hairpins). At this point, translation/libration/screw (TLS) parameters were calculated for each protein chain and for arm H of the junction, and were incorporated into the refinement. TLS parameters were also calculated for the remaining three arms of the junction, but their inclusion led to instability in the refinement and so they were subsequently omitted.

Approximately 95% of the data were used for refinement whereas the remaining 5% were used to calculate R_{free}. No electron density could be located for residues 1–15 and 146–149 of the protein or for the 3-base thymine hairpins located at the ends of two of the DNA arms. The final model displays good stereochemistry, with 92.4% of residues in most-favoured regions of the Ramachandran plot and the remaining 7.6% in additionally allowed regions as defined by Procheck[34].

30. Déclais, A. C., Liu, J., Freeman, A. D. & Lilley, D. M. Structural recognition between a four-way DNA junction and a resolving enzyme. *J. Mol. Biol.* **359,** 1261–1276 (2006).
31. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsl. Protein Crystallogr.* **26** (1992).
32. Evans, P.R. SCALA. *Joint CCP4 + ESF-EAMCB Newsl. Protein Crystallogr.* 33 22–24 (1997).
33. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60,** 432–438 (2004).
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 (1993).

©2007 Nature Publishing Group

The personal side of genomics

Innovations in DNA sequencing and genotyping are opening doors for personal genomics. Nathan Blow explores these technological advances and their implications.

The era of personal genomics is upon us, with advances in technologies such as DNA sequencing and genotyping fuelling the fires. Personal genomics is a story of researchers looking for genetic clues to our most common diseases, of dazzling advances in genetic analysis technology and of lingering questions about how the public will view and use the information.

DNA sequencing is clearly driving much of this revolution in personal genomics. In late May 2007, 454 Life Sciences in Branford, Connecticut, and the Human Genome Sequencing Center at the Baylor College of Medicine in Houston, Texas, made headlines around the world with the announcement that they had sequenced James Watson's entire genome using 454 Life Sciences' next-generation sequencing technology. And just four months later researchers at the J. Craig Venter Institute in Rockville, Maryland, along with collaborators at The Hospital for Sick Children in Toronto, Canada, and the University of California, San Diego, published the first full genome sequence of a single individual — Craig Venter[1]. This analysis, though, relied on the traditional approach of Sanger sequencing.

Now, some groups are looking to take DNA sequencing and personal genomics to even higher levels. "We want to look at 100,000 genomes and rather than just look at the genomics, in which you get an idea of variation like with the HapMap, we want to actually look at the trait connected with the variation and the environment," says George Church of Harvard Medical School in Boston, Massachusetts, and founder of the Personal Genome Project (see 'Being well informed').

M. NEMCHUK, BROAD INST.

The Broad Institute's Chad Nusbaum uses several next-generation sequencing systems in his research.

When first conceived in 2003, the Personal Genome Project faced numerous challenges, not least that the technology required to meet its goals was not even available. But technology is catching up with ambition, and advances in DNA sequencing are making it possible to decode individual genomes much faster, making endeavours such as the Personal Genome Project more feasible.

Sequencing's new wild west

A new generation of faster DNA-sequencing systems has exploded onto the genetic-analysis scene, with at least five companies offering or preparing to offer sequencers that boast amazing output. Several sequencers produce upwards of a billion bases of raw data per run — the equivalent of one-third of the human genome.

But Chad Nusbaum, co-director of the genome-biology programme at the Broad Institute in Cambridge, Massachusetts, is quick to point out that these new systems

BEING WELL INFORMED

George Church, of Harvard Medical School in Boston, Massachusetts, is working on a project that he thinks could change the landscapes of both genomics and medicine. His Personal Genome Project aims to integrate data for genomics, environment and phenotype in more than 100,000 volunteers.

G. CHURCH

George Church in his Personal Genomics Project picture with a ruler to measure facial features.

Perhaps the biggest issue for the project is how to acquire informed consent from so many participants, who will have their data become publicly available. This is something that Church hopes the Personal Genome Project will be able to address. He comments that some researchers think that this level of scale-up is not compatible with the current rules of informed consent.

From the start, participants will need to know what they are volunteering to do. "We are trying to emphasize to participants that rich holistic genetic and trait data are going to be obtained," says Church. Therefore, he says, participants must be aware of the risks and benefits of modern genetics and of the fact that modern digital information has various ways to get into the public domain.

To make certain that everyone involved clearly understands these points, the project will establish an online educational system targeted to these topics. All participants will need to pass a test to see how well they understand them. After this education and evaluation, the project will allocate participants 'set points', which will determine how much information should be released to them.

Additionally, Church says, participants will have access to data only for validated genes for which we know something about their action or disease risks or for which treatment is possible. Although that list is short now, he expects that it will grow quickly.

Still, some researchers are concerned about the risks of providing patients with genetic data. "I think there is a danger with a lot of the new and best technologies that it is tempting to provide sequence information to patients before the biological implications of those data are known," warns Bert Vogelstein of Johns Hopkins University in Baltimore, Maryland.

In August 2005 the project received approval from Harvard Medical School's institutional review board to start enrolling its first ten participants. The criteria were stringent: participants had to have at least a master's-level education in genetics or an equivalent understanding of genetics research. Church and others involved in the field were among the first to volunteer because they were thought to be the best informed to give consent.

Church thinks that now is the time to resolve the issues of informed consent because the technology has arrived to make personal genome sequencing a reality. He points out that companies such as 23andMe in Mountain View, California, and DNAdirect in San Francisco, either already offer personal genetic testing services or plan to in the future. "We have to get this in place before everything just goes crazy," says Church. N.B.

©2007 Nature Publishing Group

are in fact quite different from one another and at various stages of maturity. "The principle that we use when applying these new technologies is that there is a lot of expensive sequencing that we do with Applied Biosystem's 3730xl system and anything that we can move over to the new technologies, as long as it is effective, is bound to be cheaper." The systems available now from Roche, Illumina and Applied Biosystems do seem to be effective, as the Broad Institute and other organizations are using them for various sequencing-based applications.

Assembling the future

By the end of last year, 454 Life Sciences, which was founded in 2000 and was recently acquired by Roche, had more than 60 of its sequencing systems placed around the world. "Our technology is in all major US genome centres and some of the international centres," says Michael Egholm, vice-president of research and development at 454 Life Sciences.

The technology developed by 454 Life Sciences is based on two fundamental principles: emulsion PCR and pyrosequencing. Emulsion PCR side-steps the conventional process of bacterial cloning by attaching fragments of DNA 300 to 500 base pairs long to beads *in vitro,* then amplifying them with PCR to make millions of identical copies. Pyrosequencing allows for a massive parallel reaction format done in 1.6 million wells on a PicoTiterPlate. "Right now, day in and day out, we can perform 400,000 reads of 250 bases each with an accuracy of 99.5% or better," says Egholm. Although the 454 Life Sciences system is not as accurate as conventional Sanger sequencing, Egholm notes that it is an order of magnitude more productive (see 'Truth and accuracy').

This upgraded Genome Sequencer FLX System allows more sequencing cycles and therefore longer reads than the previous Genome Sequencer 20 System. Longer reads help in whole-genome sequencing and assembly applications. "We believe that shortly there will be many more *de novo* assembled genomes due to our technology," says Egholm. He notes that the genomes of several microorganisms have been assembled from scratch by use of 454 sequencing, and the technology has also been used to supplement Sanger sequencing on a few projects involving larger genomes.

At the Broad Institute, where researchers use two FLX systems and one Genome Sequencer 20, Nusbaum appreciates the ease of the 454 sequencing process. "It is nicer than Sanger sequencing because it is a faster and simpler process." He points out that at the Broad Institute, sequencing a bacterium can take a month with Sanger methodology, whereas with 454 technology it can be done in a week and without the high degree of clone tracking associated with Sanger sequencing.

Still, for *de novo* sequencing and to assemble larger genomes, such as those of mammals, longer paired reads — that is, two reads that are a known distance apart — will be necessary — an issue that Roche and 454 are trying to address. "Whether [454 sequencing] will work with a mammalian genome is a good question, and it is a little way off," says Nusbaum. But he optimistically notes that 454 Life Sciences has exceeded his expectations in surmounting several other technical hurdles. Egholm, however, is much true biological diversity, or just errors caused by the new sequencing technology. "We had to explore just how good the sequencing technology actually was," he says.

454 LIFE SCIENCES

The Genome Sequencer FLX, developed jointly by 454 Life Sciences and Roche Applied Sciences, is based on 454 sequencing technology.

TRUTH AND ACCURACY

Mitch Sogin, director of the Josephine Bay Paul Center for Comparative Molecular Biology and Evolution at Woods Hole Marine Biological Laboratory in Massachusetts, performs environmental sampling of nucleic-acid sequences. "Every sequence has the potential to tell us an important story," says Sogin, so highly accurate analysis techniques are needed.

But when Sogin's lab switched over from traditional Sanger-based sequencing to the next-generation sequencing system of 454 Life Sciences in Branford, Connecticut, to study these environmental samples, something strange happened. "The diversity was between ten- and a hundred-fold more divergent than we expected," recalls Sogin.

So Sogin and his colleagues needed to determine whether the unexpected findings represented

M. SOGIN

Mitch Sogin tested the accuracy of 454 sequencing.

Sogin and his colleagues did a straightforward experiment in which they resequenced more than 50 templates and cloned sequences on a 454 Life Sciences Genome Sequencer 20 that they had sequenced previously with the Sanger methodology. The work showed that the 454 system was 98% accurate if no culling was used to remove bad bases or reads[6].

However, by using a very simple set of rules, which caused somewhere between 10% and 20% of the data to be discarded, the accuracy could be pushed up to 99.75%. And discarding up to 20% of the data for this level of accuracy is a trade-off that is fine with Sogin because the latest 454 system — the Genome Sequencer FLX — can produce up to 400,000 reads per run.

Others agree that for some applications the large amount of data generated by the next-generation sequencing systems could trump the accuracy produced through Sanger sequencing. "For applications such as CHIP-sequencing you can use the 454 or Solexa 1G data even though they have lower base accuracy because you do not need it. What you need is the volume for the experiment," says Chad Nusbaum of the Broad Institute in Cambridge, Massachusetts.

Sogin now thinks that traditional sequencing methods had been underestimating the biological diversity of the environmental nucleic-acid samples. "Turns out that the diversity is coming from low-abundance nucleic-acid populations that you are not likely to encounter if you sequence only a few hundred molecules. You see these low-abundance molecules only if you sequence many tens of thousands of molecules," he says. N.B.

©2007 Nature Publishing Group

more direct in his vision for the future of 454 sequencing. "My goal is simple, I want to displace Sanger sequencing for *de novo* sequencing."

Counting games

Like the Genome Sequencer FLX system, both Illumina and Applied Biosystems have used emulsion PCR as a starting point for their next-generation sequencing systems. But from there the methods of sequencing are quite different from each other.

In January this year, Illumina, located in San Diego, California, acquired the Hayward-based firm Solexa. Solexa's key technology, previously called the Solexa 1G and now named the Genome Analyzer, is a next-generation sequencing system that can sequence the equivalent of a third of the entire human genome in a single run. The Broad Institute now uses 16 Genome Analyzers for various projects. "Any application that is counting-related is a very good one to perform using Illumina's system," says Nusbaum.

ILLUMINA

The Illumina Genome Analyzer has been used for several CHIP-sequencing projects.

Nusbaum and his colleagues, along with other groups, have already demonstrated the usefulness of the Genome Analyzer in looking at patterns of chromatin structure by using chromatin immunoprecipitation[2–5]. "It is an incredibly powerful application of the technology," says Nusbaum. By pulling down DNA bound to histones carrying specific modifications, sequencing it and mapping it back to the genome, they could map the status of chromatin across the genome and throughout development. Nusbaum adds that it is also an incredibly easy application of the technology and anticipates that Illumina instruments will be used for other applications such as transcriptional profiling or microRNA and small RNA discovery. "It is also a great way to identify polymorphisms in genomes that are not extremely different."

Applied Biosystems in Foster City, California, is rolling out its new sequencing by oligonucleotide ligation and detection, or SOLiD, system in October 2007. The target is to cover a whole human genome in one run, says Kevin McKernan, senior director of scientific operations at Applied Biosystems. McKernan says that in-house, the SOLiD system has been obtaining around a gigabase more data than their target, achieving 4 gigabases of sequence per run that aligns to the target genome, and 8 gigabases overall. McKernan thinks that with advances in the PCR process over time, this will turn into 8 gigabases of sequence that aligns to the target genome. The SOLiD system differs from other next-generation sequencing systems

CHIPPING OUT OUR DIFFERENCES

Single nucleotide polymorphism (SNP) genotyping is a method for determining genetic variation. As more and more SNPs have been identified from the genome in recent years, the power of this technique has steadily increased.

Affymetrix, located in Santa Clara, California, is one company that is taking advantage of SNP-discovery projects, such as the International Hap Map effort, to generate SNP arrays for whole-genome association studies. In May, Affymetrix launched its next-generation array, called the Genome-wide Human SNP Array 6.0, or SNP6.0. "This chip allows us to look simultaneously at more than 1.8 million markers of genetic variation," says Keith Jones, vice-president of assay and application product development at the company.

The SNP6.0 not only offers genome-wide SNP coverage, but also contains more than 900,000 probes that target copy-number variants (CNVs) in the genome. "When walking down the path in designing the SNP6.0, we took the biochemistry that we used to generate targets to hybridize to the arrays and empirically identified probes that responded in a dose-dependent manner to changes in copy number," says Jones.

Other companies have also placed CNV probes onto their genotyping chips. Illumina, based in San Diego, California, now offers the Human 1M BeadChip, which boasts more than one million SNP and CNV probes, for whole-genome genotyping applications. Nimblegen, located in Madison, Wisconsin, and recently acquired by Roche Applied Sciences, also offers several whole-genome and custom-tiling array comparative genome hybridization products for examining copy-number variation across the entire genome. These arrays contain more than 385,000 probes at a median spacing of 6,000 base pairs. Agilent Technologies, located in Santa Clara, California, provides several array comparative genome hybridization products for analysis of copy-number variation in humans, mice and rats.

AFFYMETRIX

Affymetrix now offers the Genome-wide Human SNP Array 6.0 containing more than 1.8 million markers of genetic variation.

Illumina also offers several more-focused SNP arrays, including a cancer panel and one that targets the major histocompatibility complex, an area that Affymetrix also seems to be moving into. "I think it is also fair to say that you will see more application-specific SNP panels in the future," says Jones. N.B.

©2007 Nature Publishing Group

by performing sequencing by ligation rather than by synthesis, as conventional systems do. "Traditionally, people probably don't think of ligases as being more accurate than polymerases," says McKernan. But he claims that the SOLiD system achieves its accuracy by the new way in which the probes are encoded.

By doing successive rounds of ligation and looking at particular probe colour, the SOLiD sequencer obtains information not from just one base, but also from adjacent bases. So after ligation every colour has more than one base of information, which permits multiple colour calls for each base location. By using this redundant information, McKernan says a tremendous amount of error correction can be performed when making base-call decisions. "What we have been seeing is that this is giving us a ten to twenty times improvement over polymerase-based systems in terms of raw accuracy," says McKernan.

Applied Biosystems views the SOLiD system as particularly well suited for the cancer-research community and for applications in personal genomics. "Folks in the cancer community are going to gravitate towards it because they are looking for low-frequency mutations and the higher accuracy the system delivers through our error-correction scheme will be beneficial," says McKernan. He notes that cancer genomes tend to be very complex, with copy-number changes or copy-number neutral changes, such as translocations. Using sequencing systems such as SOLiD, these changes can be visualized by watching when pair reads, or 'mate pairs', break, quickly providing information about translocations. These data are currently analysed with copy-number chip assays (See 'Chipping out our differences').

Both the Genome Analyzer and the SOLiD system have some limitations because of the short read lengths of around 35 base pairs per run. "For sequencing larger genomes, 35 bases or fewer per run just won't cut it. We tried with 100 bases for a long time and had problems," says Egholm. McKernan says that the way Applied Biosystems is attempting to resolve this issue is by using mate pairs. "Whenever there is a single read in a repetitive region you do not know where to place it," says McKernan. But he notes that with a mate pair you know that that read is linked to something 3–4 kilobases away, so those reads can be placed accurately. And this placement accuracy can be very important, as the repeat content of the human genome is high, estimated to be upwards of 50%.

APPLIED BIOSYSTEMS

The next-generation SOLiD system from Applied Biosystems uses ligation-based sequencing methods.

Pulling in the 'exome'

Church is excited by the fact that the advanced sequencing technology now in place has drastically lowered the cost of sequencing. "What has happened in the past year is that the price has plummeted by a factor of 100," he says, making a project of the scope of the Personal Genome Project realistic from a financial perspective. The tricky part now is getting the most useful information from the human genome in a similarly cost-effective manner.

ILLUMINA

Illumina now offers several SNP and copy-number-variation probe chips for genotyping applications.

For the Personal Genome Project and other groups, the challenge is to obtain just exons, or the 'exome', from the human genome. Technically, the simplest method, performing in excess of 200,000 individual PCR reactions, would also be the most labour intensive and costly. But groups of researchers and companies are now working on ways to selectively amplify or capture different parts of the genome. "Soon you will be able to cherry-pick all the way along and identify the parts of the genome that are most likely to yield the maximum information," says Church. And these methods will be very welcome additions to the genomics world because, as Church notes, not all pieces of DNA are created equal. Although he points out that no DNA is 'junk', he contends that by examining only 1% of the genome you can get about 98% of the information about positions that cause changes in traits.

The technology necessary for a personal-genomics revolution is here on the scene. Most people say that the major concern for personal-genomics projects is how to deal with the data from participants. And even on that front there seems to be a lot of optimism within the genomics community. Nusbaum is encouraged that people such as Church have taken up such initiatives as the Personal Genome Project. "I am glad that someone like George is taking this on because he has charisma and clout, so that even if people don't want to hear what he is saying, they have to listen." ■

Nathan Blow is Technology Editor for *Nature* and *Nature Methods*.

1. Levy, S. *PloS Biol.* **5,** e254 (2007).
2. Mikkelsen, T. S. *et al. Nature* **448,** 553–560 (2007).
3. Robertson, G. *et al. Nature Methods* **4,** 651–657 (2007).
4. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. *Science* **316,** 1497–1502 (2007).
5. Barski, A. *et al. Cell* **129,** 823–837 (2007).
6. Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L. & Welch, D. M. *Genome Biol.* **8,** R143 (2007).

©2007 Nature Publishing Group

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
DNA sequencing and analysis			
454 Life Sciences	Automated DNA-analysis systems including Genome Sequencer 20 and FLX	Branford, Connecticut	www.454.com
Applied Biosystems	Automated sequencing systems using both Sanger methodology and next-generation SOLiD-based system	Foster City, California	www.appliedbiosystems.com
Beckman Coulter	Automated DNA sequencing systems based on capillary electrophoresis using the Sanger method	Fullerton, California	www.beckmancoulter.com
Cogenics	Contract automated DNA sequencing services using both Sanger-based and next-generation sequencing systems	Morrisville, North Carolina	www.cogenics.com
DNAStar	Software for DNA sequence and microarray analysis	Madison, Wisconsin	www.dnastar.com
GE Healthcare	Supplies kits for automated DNA sequencing systems	Buckinghamshire, UK	www.gehealthcare.com
Geospiza	Software for management genetic data	Seattle, Washington	www.geospiza.com
Helicos Biosciences	Development of the HeliScope Genetic Analyzer based on single-molecule sequencing methodology	Cambridge, Massachusetts	www.helicosbio.com
Illumina	Illumina Genome Analyzer next-generation DNA-sequencing system	San Diego, California	www.illumina.com
Licor Biosciences	Automated DNA sequencing instrumentation using the Sanger methodology	Lincoln, Nebraska	www.licor.com
Pacific Biosciences	Developing automated next generation sequencing instrument	Menlo Park, California	www.pacificbiosciences.com
Promega	Supplies the fMol DNA sequencing system for manual DNA sequencing and promoter mapping applications	Madison, Wisconsin	www.promega.com
Roche Applied Sciences	Supplies the 454 Life Sciences Genome Sequencer 20 and FLX systems	Basel, Switzerland	www.roche.com
Saturn Biotech	Full DNA-sequencing service provider	Perth, Australia	www.saturnbiotech.com
SeqWright	Contract automated DNA-sequencing services	Houston, Texas	www.seqwright.com
Genotyping and mutation detection			
ACGT	Custom SNP genotyping and DNA-sequencing services	Wheeling, Illinois	www.acgtinc.com
Affymetrix	Suppliers of a variety of SNP and copy-number-variant chips and reagents	Santa Clara, California	www.affymetrix.com
Agencourt Bioscience Corporation	SNP discovery and DNA sequencing services	Beverly, Massachusetts	www.agencourt.com
Asper Biotech	Custom SNP assay development and genotyping applications; genetic-testing kits	Tartu, Estonia	www.asperbio.com
Biocat	Kits for array CGH and methylation analysis	Heidelberg, Germany	www.biocat.de
Bioserve	Genotyping assay development and DNA-sequencing services	Laurel, Maryland	www.bioserve.com
Biotage	SNP and methylation analysis instrumentation using pyrosequencing methods	Uppsala, Sweden	www.biotagebio.com
Bioventures	Cytochrome P polymorphism detection	Murfreesboro, Tennessee	www.bioventures.com
Bruckner Daltonics	Genolink SNP-genotyping system that uses mass spectrometry	Billerica, Massachusetts	www.bdal.com
Chemicon (Millipore)	SNP-genotyping systems for assay development	Billerica, Massachusetts	www.chemicon.com
DeCode Genetics	Contact services for human SNP and microsatellite genotyping	Reykjavik, Iceland	www.decode.com
DxS	SNP-genotyping kits for genetic testing; custom genetic analysis services including SNP genotyping, DNA isolation and sequencing	Manchester, UK	www.dxsgenotyping.com
Genes-4U	Mutations detection systems for use with the Light Cycler	Zurich, Switzerland	www.genes-4u.com
Marligen Biosciences	SNP genotyping assay kits including a genotyping kit designed especially to genotype male-specific markers on Y chromosome	Ijamsville, Maryland	www.marligen.com
Nanogen	Nanochip-automated instrumentation for SNP analysis and assay development	San Diego, California	www.nanogen.com
Nimblegen Systems	Products from methylation analysis, SNP genotyping, array CGH, and chromatin immunoprecipitation arrays.	Madison, Wisconsin	www.nimblegen.com
Orchid Cellmark	DNA-testing services using a patented SNP-based technology	Princeton, New Jersey	www.orchidbio.com
Perkin Elmer Life Sciences	SNP detection kits and reagents	Waltham, Massachusetts	las.perkinelmer.com
Sequenom	SNP discovery and detection for genotyping applications; methylation analysis	San Diego, California	www.sequenom.com
Transgenomic	Products and services for automated high-sensitivity genetic variation and mutation analysis	Omaha, Nebraska	www.transgenomic.com
Genetic and genomic testing companies			
23andMe	Newly founded company seeking to explore personal genomics	Mountain View, California	www.23andme.com
AncestryByDNA	Developing genetic-testing products and services	Sarasota, Florida	www.ancestrybydna.com
Athena Diagnostics	Diagnostic testing services for neurological disorders	Worcester, Massachusetts	www.athenadiagnostics.com
DNADirect	Genetic-testing services; genetic counselling	San Francisco, California	www.dnadirect.com
GeneDx	Genetic-testing services; mutation confirmation services	Gaithersburg, Maryland	www.genedx.com
Genetic Technologies	Genetic-testing services for people, animals and plants	Fitzroy, Australia	www.gtg.com
Genzyme Genetics	Human genetic-testing services for reproductive genetics and oncology/pathology	Cambridge, Massachusetts	www.genzymegenetics.com
IdentiGEN	Genetic testing of products for the agriculture and food industries	Dublin, Ireland	www.identigen.com

©2007 Nature Publishing Group

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL	
Gene-expression analysis				
Agilent Technologies	Microarrays; lab-on-a-chip methods; SNP analysis software	Wilmington, Delaware	www.agilent.com	
Genomic Solutions	Assay assembly for SNP genotyping; sequencing and PCR	Ann Arbor, Michigan	www.genomicsolutions.com	
Operon Biotechnologies	Oligonucleotide production; preprinted oligonucleotide arrays and probe sets	Huntsville, Alabama	www.operon.com	●
Oxford Gene Technology	Oligonucleotide arrays for various applications	Oxford, UK	www.ogt.co.uk	
QIAGEN	DNA, RNA extraction kits; genotyping; haplotyping	Germantown, Maryland	www.qiagen.com	●
TeleChem International	Microarray robots, printers, scanners and services	Sunnyvale, California	www.arrayit.com	
Vermillion	Biomarker identification for drug discovery	Fremont, California	www.ciphergen.com	
Taconic	Genotyping services for mice and rats	Germantown, New York	taconic.transnetyx.com	
Expression Analysis	Microarray-based genomic services for clinical trials	Durham, North Carolina	www.expressionanalysis.com	
General				
Active Motif	Kits and reagents for nucleic-acid isolation; kits for performing chromatin immunoprecipitation experiments	Carlsbad, California	www.activemotif.com	
Brinkmann Instruments	Laboratory instrument supplier	Westbury, New York	www.brinkmann.com	
Cambrex	Molecular and cellular biology reagents and tools	East Rutherford, New Jersey	www.cambrex.com	
Clontech	Reagents for PCR, nucleic-acid extraction and purification; RNAi expression systems	Mountain View, California	www.clontech.com	
EMD (Merck)	Reagents and kits for molecular biology	Madison, Wisconsin	www.emdbiosciences.com	
Epicentre Biotechnologies	Enzymes for PCR and RT-PCR; DNA and RNA purification and isolation kits	Madison, Wisconsin	www.epibio.com	
Fisher Scientific	Catalogue supplier of research equipment, chemicals and kits	Waltham, Massachusetts	www.fishersci.com	
Hamilton Company	Robotics for the life sciences	Reno, Nevada	www.hamiltoncomp.com	
Inivtrogen	Kits and reagents for molecular biology, genomics and cell biology research	Carlsbad, California	www.invitrogen.com	●
Lonza	Molecular biology reagents and systems	Basel, Switzerland	www.lonza.com	●
Maxim Biotech	PCR reagents and systems; nucleic-acid isolation; custom services including primer design, sample preparation and library construction	San Francisco, California	www.maximbio.com	
Merck	Chemicals, kits and reagents for molecular- and cellular-biology-related research	Darmstadt, Germany	www.merck.de	
MP Biomedicals	Reagents and chemicals for research	Aurora, Ohio	www.mpbio.com	
New England Biolabs	Molecular-biology-related reagents, kits and enzymes	Ipswich, Massachusetts	www.neb.com	
Stratagene	Tools for molecular biology, proteomics and genomics research	La Jolla, California	www.stratagene.com	
Takara Bio	Reagents, kits and services for genomics and molecular biology research	Shiga, Japan	www.takara-bio.com	●
USB	Chemicals and reagents for molecular biology	Cleveland, Ohio	www.usbcorporation.com	
ITSI Biosciences	RNA profiling and quantitation services	Johnstown, Pennsylvania	www.itsibio.com	
Open Biosystems	Gene expression products, RNAi genomic engineering and transfection reagents	Huntsville, Alabama	www.openbiosystems.com	
OriGene	Expression products, RNAi reagents, antibodies	Rockville, Maryland	www.origene.com	
GenHunter	Differential display, PCR cloning, AP-TAG technology	Nashville, Tennessee	www.genhunter.com	
NuGen	Products to enable whole transcriptome gene-expression analysis	San Carlos, California	www.nugeninc.com	

● see advertisement

©2007 Nature Publishing Group

naturejobs

JOBS OF THE WEEK

About 900 young scientists came to seek careers advice at The Source last week, an event in London sponsored by *Naturejobs*. In response, expert panellists emphasized that, as well as seeking external advice from those with experience, it is just as important to look inwards to ask yourself what kind of project, collaboration and field you want to join.

For instance, in an earlier session on alternative careers that I chaired, two panellists said that they knew before leaving graduate school that academia wasn't for them. Franck Tetaz, a partner in a French patent-protection firm, joked that it was the repeated breakage of beakers and test tubes that convinced him that he shouldn't be a chemist. Fraser Campbell also knew he wasn't meant for the tenure track when he began wondering about the real-world outcomes of research. He found them in the business world, first as a biotech analyst and now in corporate finance.

Both were quick to point out that there are parts of their jobs they don't enjoy. For Tetaz, it's dealing with an in-tray stacked high with legal documents. For Campbell, it's grappling with spreadsheets. But the upside for both is being exposed to a variety of scientific problems and looking for real-world solutions that might result in products.

So at The Source, I told the young researchers to ask themselves what they enjoy most and least about their scientific lives. Graduate students and postdocs sometimes forget that activities such as writing a grant, supervising lab workers, ordering equipment and searching the literature are as legitimate examples of doing science as performing experiments. Specializing in any one of those activities can itself provide a satisfying scientific career. Once you identify what you enjoy doing and are good at, questions about which sector, discipline and part of the world you're interested in tend to resolve themselves.

Paul Smaglik was editor of *Naturejobs* from 2001 to 2007.

Correction

In the Regions feature 'Enter the dragon' (*Nature* **449,** 502–504; 2007), we quoted Yi Yinfa as saying that the curriculum at the Shenzhen Virtual University Park is determined in a "market oriented" fashion. This quote should in fact have been attributed to the university park's director, Qiu Xuan.

CONTACTS
Acting Editor: Gene Russo

US Head Office
New York
75 Varick Street,
9th Floor,
New York,
New York 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager/Corporations:
Peter Bless
Tel: +1 800 989 7718

San Francisco Office
Classified Sales Representative:
Michaela Bjorkman
West USA/West Corp. Canada
225 Bush Street, Suite 1453
San Francisco,
California 94104
Tel: +1 415 781 3803
Fax: +1 415 781 3805
e-mail: m.bjorkman@naturesf.com

European Head Office
London
The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

***Naturejobs* web development:**
Tom Hancock

***Naturejobs* online production:**
Jasmine Myer

Japan Head Office
Tokyo
Chiyoda Building,
2-37 Ichigayatamachi,
Shinjuku-ku,
Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
Tel: +81-3-3267-8765
e-mail: a.watanabe@natureasia.com

©2007 Nature Publishing Group

NATURE|Vol 449|4 October 2007

MOVERS

Ren Wang, director, Consultative Group on International Agricultural Research, Washington DC

2000–07: Deputy director-general for research, International Rice Research Institute, Manila, Philippines
1994–2000: Vice-president, Chinese Academy of Agricultural Sciences, Beijing, China
1993–95: Deputy director of programme development, International Institute of Biological Control, Ascot, UK

Ren Wang understands the importance of seizing an opportunity. As a farm worker during China's Cultural Revolution, he was given two choices: either go to an agricultural college or work in construction. His choice landed him in the plant-protection research division at Shanxi Agricultural University in Taigu.

A postgraduate at the Chinese Academy of Agricultural Sciences (CAAS) Institute of Biological Control in Beijing, Wang studied the use of beneficial insects. A PhD fellowship from the Rockefeller Foundation then took him to Virginia Polytechnic Institute in Blacksburg, where he became an expert in the control of exotic weeds.

After returning to the CAAS to continue studying the use of insects in weed control, Wang continued US–China collaborations, eventually helping to set up a joint biological-control research institute. Soon afterwards, Chinese policy moved towards commercializing research.

"I didn't want to commercialize rice, so I looked for opportunities to continue pure research," says Wang. Asked by the director of Britain's International Institute of Biological Control for personnel recommendations, Wang offered himself and became its deputy director.

Lured back to China as vice-president of the CAAS, Wang supervised 10,000 staff. One of his major projects was a Sino–Japan Center for Sustainable Agriculture.

Moving to the Consultative Group on International Agricultural Research (CGIAR), an international public–private partnership that oversees 15 research centres, Wang became deputy director-general for research at the CGIAR's International Rice Research Institute (IRRI) in the Philippines. There he fostered collaborative initiatives between sub-Saharan Africa and central Asia.

After seven years, Wang has now become the CGIAR's director. He plans to address challenges such as climate change, poverty in Africa and South Asia and the effect of biofuel production on small farms. "If we can build strong collaborations, we'll form an unprecedented force to address such continued, pressing issues," says Wang.

His colleagues value his commitment to strengthening the CGIAR's research institutions. Robert Zeigler, director-general of the IRRI, says Wang understands the importance of decent infrastructure, a steady stream of exchanges and a critical mass of resources to address complex problems.

"I am an enabler," says Wang. "I want to develop the infrastructure that enables scientists to develop innovations." ■

Virginia Gewin

NETWORKS & SUPPORT

Careers on the move

Do European scientists need to be mobile in order to have successful careers? The Young European Biotech Network's Careers in Life Sciences (CiLS) working group conducted a study to investigate why fledgling scientists considered moving abroad. The European Job Migration Project began in 2005 with a survey of 900 students, professors and industry professionals from 70 countries (see *Nature* **435,** 709; 2005), continuing in 2006 with 20 follow-up testimonials about aspirations and attitudes.

Our analyses indicate a substantial 'brain circulation' (temporary international moves) in the European life sciences and help to identify drivers of international mobility. People planning to move to another country place importance on living conditions and total compensation packages, rather than on high salaries alone. Results from 2005 suggest that those most interested in going abroad want either to move to Britain or the United States, or are interested only in moving within continental Europe. The 2006 data suggest a connection between the wish to improve language skills and the country of choice; about 70% of those wanting to move to Britain or the United States were not native English speakers. People who have had previous experience abroad (especially internships) have greater interest in furthering their career internationally. According to the 2005 survey, men and women have roughly equal interest in working abroad; but men take longer stays. In the 2006 study, women mention family reasons as a motivator to stay in or return to their home country. People recruited by job agencies go abroad significantly less frequently than those who find work through newspapers or the Internet.

These results have provided a new direction for the CiLS. We are initiating a survey to find out what skills young scientists consider valuable. We will also interview life-science companies, such as drug firms, consultancies, biotechs and financial institutions. We'll examine how often scientists see moves abroad as permanent, and the perceived value of moves in terms of new skills, knowledge and contacts. We'll investigate how the amount of time spent abroad and the stage of career affect people's plans. Given the extensive, contentious discussions about career mobility in Europe, we hope the project will help to identify key factors driving and impeding European scientific careers. ■

Emilia Danilowicz is a member of the Careers in Life Sciences project team.

www.yebn.org

POSTDOC JOURNAL

Small advantages

Since returning to work after having a child, I have been thinking a lot about what I will do after my fellowship. I am lucky to have friends in different career paths that I can talk to.

One of my friends had a baby the same week I did, and I asked her about how she will adjust in her job as a new mother. She is a tenured professor in environmental studies at a small private college. It offers flexibility, supportive colleagues and positive student interactions. Of course she faces some challenges, such as trying to do meaningful and successful research projects with undergraduates. At a small college one can't be too specialized, so she mentors thesis projects in a wide range of subjects, not necessarily in her area of expertise. Teaching institutions require a lot of grading, so over-scheduling is common. But hers is very family friendly, allowing her to take a full year off and work part-time when she returns.

For some, small private colleges are a perfect solution to the 'research versus teaching' quandary, as they provide healthy doses of both. Nice towns, small schools, long-term faculty and teaching opportunities await those who choose this path. It seems that small colleges afford you greater flexibility as long as you can remain flexible in your goals as well. It's good to know such opportunities exist, although I am still not sure I will follow that route. ■

Moira Sheehan is a postdoc in the Department of Plant Breeding and Genetics at Cornell University.

©2007 Nature Publishing Group

natu-oct-04-07-p000 29/9/07 7:32 Page 1

Ateliers de formation Inserm
101 rue de Tolbiac
75654 Paris Cedex 13 France

Tel: 33 (0) 144.23.62.03
Fax: 33 (0) 144.23.62.93
ateliers@tolbiac.inserm.fr

Inserm

French Institute of Health and Medical Research

Ateliers de formation 2008

182 – Biophysical modeling and mathematical analysis in cellular biology
Organisers: Thierry Galli *(Inserm/Institut Jacques Monod, Paris)*, David Holcman *(Ecole Normale Supérieure, Paris)*.

• Phase I - Critical assessment: March 13-14, 2008
Maximum number of participants: 80.
Tentative list of speakers: Robert Austin *(Princeton, USA)*, Daniel Choquet *(France)*, Maïté Coppey-Moisan *(France)*, Nicolas Destainville *(France)*, Michael Elbaum *(Israel)*, Bart Ermentrout *(Pittsburgh, USA)*, Thierry Galli *(France)*, Ari Helenius *(Switzerland)*, David Holcman *(France)*, Frank Jülicher *(Germany)*, Pierre-François Lenne *(France)*, Samir Mitragotri *(Santa Barbara, USA)*, Gary Nolan *(Stanford, USA)*, Laurence Salomé *(France)*, Stanislav Schvartsman *(Princeton, USA)*.

• Phase II - Technical workshop: November 2008: Paris/Toulouse
Training group: 18
Registration deadline: January 11, 2008

183 – Molecular interactions: theory and biophysical tools for quantitative measurements
Organisers: Danièle Altschuh *(UMR CNRS/ULP 7175, Illkirch)*, Pierre Bongrand *(UMR Inserm U600/CNRS 6212, Marseille)*.

• Phase I - Critical assessment: April 10-11, 2008
Maximum number of participants: 80.
Tentative list of speakers: Alain Brisson *(France)*, Gregory de Crescenzo *(Montréal, Canada)*, Hélène Feracci *(France)*, Peter Hinterdorfer *(Austria)*, Steve Homans *(UK)*, Deborah Leckband *(Urbana, USA)*, Eric Perez *(France)*, Jacob Pielher *(Germany)*, Frédéric Pincet *(France)*, Pierre-Henri Puech *(France)*, Pierre Schaaf *(France)*, Gideon Schreiber *(Israel)*, Olivier Thoumine *(France)*, Cheng Zhu *(Atlanta, USA)*.

• Phase II - Technical workshop: May-June and September-October 2008: Bordeaux/Marseille/Paris/Strasbourg
Training group: 2-6 per laboratory
Registration deadline: February 8, 2008

184 – Statistical Modeling and Analysis of Biological Networks
Organisers: Emmanuel Barillot *(Institut Curie, Paris)*, Stéphane Robin *(AgroParisTech/INRA UMR0518, Paris)*, Jean-Philippe Vert *(Ecole des Mines, Paris)*.

• Phase I - Critical assessment: May 15-16, 2008
Maximum number of participants: 80.
Tentative list of speakers: Uri Alon *(Israel)*, Albert-Laszlo Barabasi *(Notre Dame, USA)*, Jacq Bernard *(France)*, Florence d'Alché Buc *(France)*, Hidde de Jong *(France)*, Hervé Isambert *(France)*, Alain Lilienbaum *(France)*, Marceline Kauffman *(Belgium)*, Eugène Novikov *(France)*, Ovidiu Radulescu *(France)*, Thomas René *(Belgium)*, Denis Thieffry *(France)*, Marc Vidal *(Boston, USA)*.

• Phase II - Technical workshop: 2nd half of 2008: Paris
Training group: 20
Registration deadline: March 14, 2008

185 – Intrinsically disordered proteins and associated diseases: prediction, characterization and function
Organisers: Richard Haser *(IBCP, Lyon)*, Sonia Longhi *(AFMB, Marseille)*.

• Phase I - Critical assessment: May 19-20, 2008
Maximum number of participants: 80.
Tentative list of speakers: Martin Blackledge *(France)*, Keith Dunker *(Indianapolis, USA)*, Jane Dyson *(San Diego, USA)*, David Eliezer *(New York, USA)*, Patrice Gouet *(France)*, Sonia Longhi *(France)*, Richard Kriwacki *(Memphis, USA)*, Guy Lippens *(France)*, Anne Poupon *(France)*, Véronique Receveur-Bréchot *(France)*, Peter Tompa *(Hungary)*, Vladimir Uversky *(Indianapolis, USA)*.

• Phase II - Technical workshop: June 2008: Lyon
Training group: 10
Registration deadline: March 14, 2008

186 – Functional in vivo imaging methods: from molecules to cells
Organisers: Angela Giangrande, Michel Labouesse *(IGBMC, Strasbourg)*.

• Phase I - Critical assessment: June 5-6, 2008
Maximum number of participants: 80.
Tentative list of speakers: Jérôme Collignon *(France)*, Magdalena Goetz *(Germany)*, Stephen Hell *(Germany)*, Thomas Lecuit *(France)*, Erez Raz *(Germany)*, William Schafer *(UK)*, François Schweisguth *(France)*, Daniel St Johnston *(UK)*, Antoine Triller *(France)*, Cornelius Weijer *(UK)*, Joachim Wittbrodt *(Germany)*.

• Phase II - Technical workshop: September 15-19 2008: Paris/Strasbourg
Training group: 3-4 per laboratory
Registration deadline: April 4, 2008

187 – Analysis of DNA replication and genomic instability of DNA combing and other single-molecule assays
Organisers: Jean-Sébastien Hoffmann *(IPBS, Toulouse)*, Philippe Pasero *(IGH, Montpellier)*, Étienne Schwob *(IGMM, Montpellier)*.

• Phase I - Critical assessment: June 26-27, 2008
Maximum number of participants: 80.
Tentative list of speakers: Aaron Bensimon *(France)*, Michelle Debatisse *(France)*, Olivier Hyrien *(France)*, Dean Jackson *(UK)*, Massimo Lopes *(Switzerland)*, Raymond Monnat Jr *(Seattle, USA)*, Paolo Norio *(New York, USA)*, Marie-Jeanne Pillaire *(France)*, Christophe Place *(France)*, Nick Rhind *(Worcester, USA)*, Jean-Louis Viovy *(France)*.

• Phase II - Technical workshop: October 2008: Montpellier/Toulouse
Training group: 4 per laboratory
Registration deadline: April 11, 2008

188 – Axonal transport defects and neurodegenerative diseases: strategies to understand and treat these pathologies
Organisers: Thierry Galli *(Inserm/Institut Jacques Monod, Paris)*, Frédéric Saudou *(Institut Curie, Orsay)*.

• Phase I - Critical assessment: September 18-19, 2008
Maximum number of participants: 80.
Tentative list of speakers: Patricia Bassereau *(France)*, Fabrice Cordelières *(France)*, Carlos Dotti *(Belgium)*, Alexandra Dürr *(France)*, Thierry Galli *(France)*, Sandrine Humbert *(France)*, Erika Holzbaur *(Philadelphia, USA)*, John Kendrick-Jones *(UK)*, Judith Melki *(Israel)*, Chris Miller *(UK)*, Frédéric Saudou *(France)*, Giampietro Schiavo *(UK)*, Tom Schwarz *(Boston, USA)*, Michael Sendner *(Germany)*, Kristen Verhey *(Ann Arbor, USA)*.

• Phase II - Technical workshop: October/November: Paris
Training group: 6-8 per laboratory
Registration deadline: July 18, 2008

189 – *In vivo* and *in vitro* chemical mutagenesis in the mouse: recent progresses and applications for high-throughput identification of phenotypic or allelic series
Organisers: Philip Avner *(Institut Pasteur, Paris)*, Véronique Blanquet *(UGMA, Limoges)*, Yann Hérault *(Institut de Transgénose, Orléans)*.

• Phase I - Critical assessment: October 2-3, 2008
Maximum number of participants: 80.
Tentative list of speakers: Johan Auwerx *(France)*, Philip Avner *(France)*, David Beier *(Boston, USA)*, Véronique Blanquet *(France)*, Steve Brown *(UK)*, Sophie Chantalat *(France)*, Roger Cox *(UK)*, Martin Hrabe de Angelis *(Germany)*, Ian Jackson *(UK)*, Monica Justice *(Houston, USA)*, George Kollias *(Greece)*, John T. Kung *(Taiwan)*, Juergen Laufs *(Germany)*, Terry R. Magnuson *(Chapel Hill, USA)*, John C. Schimenti *(Ithaca, USA)*, Karen L. Svenson *(Bar Harbor, USA)*.

• Phase II - Technical workshop: November-December 2008: Évry/Limoges/Orléans
Training group: 4-6 per laboratory
Registration deadline: August 1, 2008

190 – TLR receptors: from research to medical applications
Organisers: Charles Hétru *(IBMC, Strasbourg)*, Valérie Quesniaux *(Institut de Transgénose, Orléans)*.

• Phase I - Critical assessment: October 16-17, 2008
Maximum number of participants: 80.
Tentative list of speakers: Julie Magarian Blander *(New York, USA)*, Dominique Butoni-Gatel *(France)*, Marco Colonna *(St Louis, USA)*, Isabelle Couillin *(France)*, Brian Foxwell *(UK)*, Nick Gay *(Cambridge, UK)*, François Huaux *(Belgium)*, Jean-Luc Imler *(France)*, Jean-Paul Mira *(France)*, Jane A. Mitchell *(UK)*, Muriel Moser *(Belgium)*, Luke O'Neill *(Ireland)*, Valérie Quesniaux *(France)*, Jean-Claude Sirard *(France)*, Mustapha Si-Tahar *(France)*, Patrick Squiban *(France)*, Philip Taylor *(UK)*, François Trottein *(France)*.

• Phase II - Technical workshop: 19-21 November 2008: Orléans
Training group: 8
Registration deadline: July 10, 2008

191 – Small animal imaging: medical techniques for in vivo anatomical, functional and molecular imaging
Organisers: Sylvie Chalon *(Inserm U619, Tours)*, Marc Janier *(Animage, Bron)*, Chantal Rémy *(Grenoble Institut des Neurosciences, Grenoble)*.

• Phase I - Critical assessment: October 20-21, 2008
Maximum number of participants: 80.
Tentative list of speakers: Claire Billotey *(France)*, Irène Buvat *(France)*, Jean-Luc Coll *(France)*, Patrick Cozzone *(France)*, Cyril Desvignes *(France)*, Philippe Hantraye *(France)*, Olga Millan *(Spain)*, Laurent Monassier *(Illkirch,France)*, François Moutou *(France)*, Bertrand Tavitian *(France)*, Annemie van der Linden *(Belgium)*, Laurence Vico *(France)*.

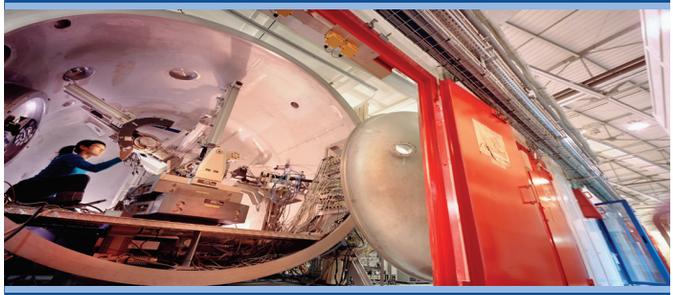
• Phase II - Technical workshop: October-November 2008: Grenoble/Lyon/Marseille/Orsay/Tours
Training group: 8-10 per laboratory
Registration deadline: August 8, 2008

W112154E

ESRF

European Sychrotron Radiation Facility

We Highlight Science



Physics and Chemistry, Life Sciences and Medicine, Earth and Environmental Sciences, Surface and Materials Sciences.

The European Synchrotron Radiation Facility (ESRF) is Europe's most powerful light source. The ESRF offers you an exciting opportunity to work with international teams using synchrotron light in Grenoble, in the heart of the French Alps.

Have a look at our vacancies at
www.esrf.eu/jobs

Contact us at **recruitment@esrf.eu**

Scientists - Post doctoral fellows - PhDstudents - Engineers - Technicians - Administrative staff

European Synchrotron Radiation Facility
ESRF, BP 220, F-38043 Grenoble Cedex 9, FRANCE,
Tel.+33 476 88 20 00 www.esrf.eu

W113787R

The Hospital of La Salpêtrière was founded in 1656 as a hospice to contain beggars, elderly, cripples and loose women. During the XVIIIth century, Philippe Pinel made significant contributions towards improving the management of psychiatric patients at La Salpêtrière. It is at the end of the XIXth century that La Salpêtrière became the birthplace of Clinical Neurology, with the tremendous contribution of J.-M.Charcot and his students including J. Babinski, G. Gilles de la Tourette, P. Marie, D. Bourneville, etc...

Nowadays, the "Groupe Hospitalier Pitié-Salpêtrière" is a general hospital with one of the most prestigious medical schools in France, and it has maintained a strong tradition in Neurology and Neuroscience. Research in Neurosciences is grouped into the Institut Fédératif des Neurosciences (IFRNS or IFR70). The plan to have in the very near future (end of 2009-beginning of 2010) a new building on the campus entirely devoted to research in neurosciences has been instrumental in shaping our decision to create a Research Centre: Le Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière (CRICM).

In our research centre, the basic unit, the brick on which the centre is built, is the team. As the specificity of La Salpêtrière is that our research is in permanent contact with, and stimulated by an amazingly abundant recruitment of neurological patients (about 100 000 per year), the richest in the country, our activities are characterized by the cross fertilization and an harmonious balance between basic and clinical (or applied) research.

So far, 33 teams have joined our project and form a matrix of 5 thematic axes: i) "Neurodegeneration", coordinated by Etienne Hirsch and Bruno Dubois; ii) "Excitability, synapse, and associated pathologies", coordinated by Richard Miles and Bertrand Fontaine; iii) "Development, glial pathology, and repair", coordinated by Catherine Lubetzki and Anne Baron-Van Evercooren; iv) "Cognition – Emotion - Action", coordinated by Rose Katz and Roland Jouvent; v) "Pains", coordinated by Jean Jacques-Benoliel. Orthogonal to these axes are teams, which aim is to develop strong methods and innovative technologies mostly in the fields of signal processing for MEG-MRI and biotechnologies.

The teams have access to a series of technical platforms. Some are common to the campus: Centre for Functional Exploration (CEF) for animal housing and experimental behavioural testing; Pitié-Salpêtrière Post-Genomic Platform (P3S) for genomic and post-genomic investigation; Pitié-Salpêtrière Cellular Imaging Platform (PIC-PS) for cellular imaging, cytometry for cell analysis and sorting. Others are more specific to the neurosciences: MEG/EEG; Centre for MRI and fMRI (CENIR); Gait, Equilibrium, Posture, and Movement; Centre for Cognitive Anatomy (CAC); DNA and Cell Bank; CNS Tumor Bank.

http://www.ifrns.chups.jussieu.fr/

UNIVERSITÉ PIERRE & MARIE CURIE

W113973R

OPPORTUNITY AT
THE UNIVERSITY OF GENEVA

THE FACULTY OF MEDICINE of GENEVA is seeking applications for a position of:

FULL or ASSOCIATE or ASSISTANT PROFESSOR OF MEDICAL BACTERIOLOGY

In the Department of microbiology and molecular medicine

This position involves full-time responsibilities for teaching at the graduate and post graduate level medical bacteriology, in particular cellular or molecular microbiology.

Candidates should have a good practical experience in the field and should be willing to participate in interdisciplinary projects, should be able to develop and direct research programs in a particular aspect of the field and to coordinate research projects in collaboration with other medical specialities, as well as assuming all pertinent administrative tasks.

A doctorate of medicine (MD), a PhD in sciences or an equivalent degree is required, as is a good knowledge of the French.

The starting date for the position is March 1st, 2008, or according to agreement.

Information concerning applications and job description are available from Stephane.jouve@medecine.unige.ch - Tel: +41 22 379 50 05 – Fax: +41 22 379 50 02

Applications must be sent before the November 29th, 2007, to:
The Dean of the Faculty of Medicine
Centre Médical Universitaire
1 rue Michel-Servet
CH-1211 Genève 4
Switzerland

Women are encouraged to apply

UNIVERSITÉ DE GENÈVE
FACULTÉ DE MÉDECINE

W113476R

ENP
École des Neurosciences Paris Île-de-France

Neuroscience

The Paris School of Neuroscience (ENP) is a network of excellent laboratories in the Paris area, within major universities and research institutes. ENP aims at facilitating graduate training and promoting outstanding research. ENP scope covers all areas of Neuroscience and associated methodologies, from fundamental to clinical and from molecular to cognitive.

ENP offers:
- Doctoral training: Applications in September and March,
- Positions for outstanding senior and junior group leaders and post-doctoral fellows,
- Summer Schools and thematic international meetings (schedule to be announced in 2008).

Founding institutions:
CEA, CNRS, Inserm, Université Pierre & Marie Curie (Paris-6), Université Paris-Sud 11.

Host institutes:
Collège de France, École normale supérieure, ICM, Alfred Fessard and Fer à Moulin Institutes, Pasteur Institute, Salpêtrière Hospital, Neurospin, Campuses of Universities Paris 5, 6, 11, and other institutions.

For further information:
http://www.enp.paris-neuroscience.fr

W112682R

natu-oct-04-07-p000 1/10/07 9:49 Page 3

Institut de Myologie

Website: http://www.institut-myologie.org
contact@institut-myologie.org

The Institute of Myology was created in 1996 as a novel structure where medical doctors of different disciplines and researchers were grouped around the patients with neuromuscular disease. This idea of public-private partnership was initiated by the French Muscular Dystrophy Association –(Association Française contre les Myopathies - AFM) and was endorsed by the University Pierre et Marie Curie (Paris VI), the Hospital of the Pitié-Salpêtrière (AP-HP), the Inserm, the CNRS and the CEA. Thomas Voit is its Medical and Scientific Director since 2006.

The main objective was to create a **translational centre of expertise on muscle research and muscle disease** -from fundamental research to patient care- by providing cutting edge research, high clinical expertise, teaching of Myology and performing technical platforms.

The Institute has since developed into the largest **French Neuromuscular Reference Centre** with more than 5.000 patients per year. It proposes a global approach to the patients and their families including diagnosis, treatment, genetic counselling, and regular follow-up in an interdisciplinary setting.

The Institute is also a first class research structure (around 200 persons involved) which facilitates translational research from the bench to the bed of patient :

- **Two research units** conduct programs on myogenesis, muscle genetics, physiopathology, and gene and/or cell therapy;

- **The NMR Laboratory and the Neuromuscular Physiology and Evaluation Laboratory** develop methodologies and tools to improve means dedicated to muscle evaluation;

- **The Morphology Unit** provides identification and morphological characterisation of paediatric and adult neuromuscular disease;

- **The Tissue Bank for Research**, member of Eurobiobank, allows easy access to large tissue collections for research purposes.

All the teams of the Institute are involved in development and/or implementation of preclinical and clinical trials and the Institute is **dedicated to the development of innovative modes of therapy, including enzyme replacement, cell and gene biotherapies, for neuromuscular disease**. This includes preclinical studies in the large animal. The clinical protocols also cover natural history studies and Phase I, II and III pharmacotherapy trials.

Teaching activities include information, training, and doctoral schools intended for all disciplines interacting with Myology, and offers an inter-university diploma in Myology. The objective is to encourage dissemination of knowledge about striated muscle and its diseases and thus to contribute to the recognition of Myology in the medical and scientific field.

The Institute offers medical and research positions –from several months to years- to foster cross-fertilization with other national and international centres.

The Laboratory Traffic and signalisation in epithelial cells (Trafic et signalisation dans les cellules épithéliales, INSERM UMR 538 University Pierre et Marie Curie), Paris, France, invites applications for a

Post-doctoral position
Peptide-membrane interactions (one or two years)

The successful candidate will participate in a research program on the investigation of the molecular relationship between peptides and lipid membranes. This research program concerns cell penetrating peptides and antibacterian peptides. The research is the core of an ANR-PCV program grant of Prof. Germain Trugnan that aims to define the role of lipid organization within membrane microdomains on the penetration of peptide vectors and on the action of antimicrobial peptides. The research approach encompasses a powerful combination of biochemical, biophysical and molecular approaches.

The candidate will be enthusiastic with a background in physicochemical disciplines. The candidate should have a Ph.D. degree or equivalent. A comprehensive knowledge of membrane cell biology is appreciated and affinity with biophysical methods is an advantage.

We offer an experienced, collaborative and international group of colleagues and an excellent infrastructure. We take part of a consortium (and a future Institute) of multidisciplinary research and are engaged in many collaborations. This ensures a wide spectrum of research approaches and facilities. The salary will be in accordance with the French salary scales for academics.

Additional information can be obtained by e-mail (Germain Trugnan: trugnan@ccr.jussieu.fr **and Jesus Ayala-Sanmartin:** jayala@chusa.jussieu.fr). and visiting http://perso.orange.fr/groupe.nj.conte

Applicants should send their application and CV including 2 references to G. Trugnan or J. Ayala-Sanmartin. Support letters are welcome. Mailing address: INSERM U538, Trafic et signalisation dans les cellules épithéliales, CHU Saint Antoine, 27 rue Chaligny, 75012 France. Tel: 33+1 40 01 13 23 and 24; Fax: 33+1 40 01 13 90.

Closing date: October 30, 2007

UNIVERSITE PIERRE & MARIE CURIE
LA SCIENCE A PARIS

W113894R

Université d'Ottawa

University of Ottawa

Au cœur de la découverte depuis 1848

L'une des cinq premières universités de recherche au Canada, l'Université d'Ottawa attire des chercheurs du monde entier grâce à ses installations hors pair et à son engagement sans égal envers la créativité et l'innovation.

Making Discoveries Since 1848

As one of Canada's top five research-intensive universities, the University of Ottawa attracts talented researchers from around the world thanks to outstanding facilities and an unparalleled commitment to creativity and innovation.

NW113766R

uOttawa
L'Université canadienne
Canada's university

Joignez-vous à notre équipe bilingue au cœur de la capitale nationale

www.uOttawa.ca

Join our bilingual team in the heart of the national capital

www.uOttawa.ca

IN 2008
CNRS IS RECRUITING

MORE THAN 400 TENURED RESEARCHERS IN ALL FIELDS OF SCIENCE

- MATHEMATICS • PHYSICS • NUCLEAR AND HIGH-ENERGY PHYSICS
- CHEMISTRY • ENGINEERING
- SCIENCE OF COMMUNICATION AND INFORMATION TECHNOLOGY
- ASTRONOMY AND EARTH SCIENCES
- ENVIRONMENT AND SUSTAINABLE DEVELOPMENT
- LIFE SCIENCES • HUMANITIES AND SOCIAL SCIENCES

CNRS encourages junior and senior scientists from around the world to apply for its tenured researcher positions.

CNRS provides an enriching scientific environment:

- numerous large-scale facilities
- highly-skilled technical support
- multiple international and interdisciplinary networks
- access to university research and teaching
- lab-to-lab and international mobility

APPLICATION FORMS AND FURTHER INFORMATION WILL BE AVAILABLE ONLINE AT WWW.CNRS.FR IN DECEMBER 2007.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

W113679R

Institut Cochin

PARIS DESCARTES · CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE · Inserm Institut national de la santé et de la recherche médicale

Tenure-Track Positions In Vaccine Research

The Cochin Institute, National Institute of Heath and Medical Research, Paris, France

The Cochin Institute is seeking applications for one junior and one senior scientists interested by the development of a dynamic and original research team on vaccines, focused on "Induction of immune responses and vaccine design".

Purposes of investigation would include in particular:

The Cochin Institute http://www.cochin.inserm.fr is a large Biomedical Research Centre localized on the campus of one of the major hospitals in Paris, with a staff of about 600 scientists, clinicians, technicians, post-docs and students. It is organized in 46 individual teams, with special interest in infectious diseases and host-pathogen interactions, more particularly HIV. Research activities in the Institute benefit from efficient technical core facilities, including cell imaging, transcriptomics and proteomics, large animal facilities and functional genomics.

Research teams working in Immunology are currently involved in basic research, deciphering the cellular and molecular pathways leading to T cell responses in human and murine models, as well as in therapeutic applications for infectious diseases, cancers and autoimmunity. Several of them are strongly involved in translational research and clinical studies.

The new team will be closely interacting with these groups. It will also collaborate directly with the Clinical Investigation Centre of the hospital and its dedicated immunomonitoring platform, both specialized in vaccine trials.

Depending on her/his background, the selected candidates will receive administrative assistance for applying for grants and permanent positions at University Paris Descartes, CNRS (ATIP (http://www.cnrs.fr/)) and Inserm (Junior contracts, Avenir program, Join Inserm program, Interface contracts (http://www.inserm.fr/fr/)) as well as the Chairs of Excellency program and the European Mobility Program (ERC, Marie Curie).

Applicants should provide a curriculum vitae, a list of publications, a two-page summary of their research accomplishments and a brief statement of their future research goals.

They should also arrange for three letters of recommendation to be sent to:

Dr Axel KAHN, Director of the Cochin Institute and of the Alfred Jost Federative Research Institute
22, rue Méchain, 75014 Paris, France
Tel: (33) 1 40 51 64 57 Fax: (33) 1 40 51 64 73
Email: kahn@cochin.inserm.fr

W113818R

Inserm Institut national de la santé et de la recherche médicale · I N B · Génomique Fonctionnelle · Université Victor Segalen Bordeaux 2 · CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Full Professor in Bio-imaging – Neuroscience

The University of Bordeaux is seeking a highly motivated candidate group leader to create a research team in the field of cellular/tissular bio-imaging in Neuroscience with a full professor tenure position.

The position will be accompanied by funding support for running costs/equipment and a lecturer position. The research team will be affiliated to the INSERM (the French National Institute of Health and Medical Research) Research Center in Neuroscience F. Magendie (U862) and will work closely with the Center for Functional Genomics.

The Bordeaux University totalises over 40 research groups in all fields of neuroscience from Nano-Imaging of receptor trafficking to cognition. The INSERM Magendie Neuroscience Institute is part of a federation of five Neuroscience Departments (www.inb.u-bordeaux2.fr). The Center for Functional Genomics (www.pgfb.u-bordeaux2.fr) is a joint venture of seven core facilities that include optical and electron microscopy (www.picin.u-bordeaux2.fr http://www.pgfb.u-bordeaux2.fr/electronique.htm).

The successful candidate is expected to develop research in the field of Imaging for Neuroscience, either at the cellular/molecular or in vivo imaging level to investigate physiological or pathological properties of the brain. The candidate should possess strong background in optical or electron microscopy imaging and a good publication record. Previous experience in the field of Neuroscience is an advantage but not a prerequisite.

The candidate will participate in the teaching program in Neuroscience and will be involved in the creation of a new curriculum in Bio-Imaging.

CVs, statement of research interests (three pages max) and three letters of recommendation should be sent to BioImaging@u-bordeaux2.fr Deadline for applications, October 10, 2007.

W111232R

The URGV, in the South of Paris, is a recently established unit specialized in Plant Genomics supported by INRA, CNRS and Evry University.

URGV

www.versailles.inra.fr/urgv/

2 postdoctoral positions in Arabidopsis Functional Genomics

Genetic analysis of *S. typhimurium* infection of *Arabidopsis thaliana*

We have recently identified the human endopathogen *Salmonella typhimurium* to function as a pathogen on Arabidopsis. A 3 year post-doctoral position is available to identify Arabidopsis mutants with enhanced or decreased disease symptoms upon Salmonella infection.

Protein-protein interactions in *Arabidopsis thaliana*

In the context of AGRON-OMICS, a FP6 project, we are systematically looking for protein-protein interactions between growth-related proteins. Yeast 2-hybrid and in planta methods will be used. A 2 year position is available in our group with long period stays in Vidal's lab in Boston.

To apply send CV and 2 reference letters to:

Heribert Hirt: hirt@evry.inra.fr Claire Lurin: lurin@evry.inra.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE · INRA · UNIVERSITÉ D'EVRY VAL D'ESSONNE

W113784R

"Through our posting for a post-doc position at NatureJobs we have received messages from a surprisingly large number of highly qualified investigators from all over the world, and have been able to recruit a suitable candidate. Thank you very much."

Alberto Sánchez-Fueyo, MD,
Hospital Clínic Barcelona/IDIBAPS, Spain

natu-oct-04-07-p000 1/10/07 16:50 Page 5

ENGAGING

Career articles from Naturejobs

- **Naturejobs Prospect**: quick takes on career implications of current events
- **Naturejobs Special Report**: examinations of jobs issues on both sides of the bench
- **Naturejobs Careers & Recruitment**: discipline-by-discipline exploration of opportunities
- **Naturejobs Regions**: tours of scientific hubs
- **Naturejobs Movers**: traffic reports that follow high-profile scientific globe-trotters and sector-hoppers
- **Naturejobs Postdocs & Students**: guides to taking a step toward a permanent position

www.naturejobs.com

naturejobs
making science work

nature publishing group npg

Université catholique de Louvain
Department of Biology
Biodiversity Research Center
Institute of Life Sciences

Full-time Tenure-track Faculty Position in Population Biology
Ref: *SC/BIOL17/2008/842*

A full-time tenure-track faculty position is available in the Unit of Ecology and Biogeography (Biodiversity Research Center), beginning in September 2008. The field of research of the new faculty member should be complementary to the research currently carried out in the Unit (http://www.ecol.ucl.ac.be/index_en.html). Possible fields of interest include conservation biology, molecular ecology, eco-physiology and evolutionary ecology, but high-profile candidates from other fields will also be considered. She/He will develop a strong research programme and supervise the work of MSc and PhD students.

Full-time Tenure-track Faculty Position in Animal Cell and Molecular Biology
Ref: *SC/BIOL13/2008/841*

A full-time tenure-track faculty position is available in the Institute of Life Sciences, beginning in September 2008. She/He will develop preferently research programmes in animal cell and molecular biology in collaboration with members of the Institute (http://www.uclouvain.be/en-isv.html) and supervise the work of MSc and PhD students.

For these positions, applicants should have a PhD or equivalent, post-doctoral experience and a strong record of international publications. Moreover, experience in and commitment to teaching at the university level and the capability to undertake scientific research at a high level and to lead a research team are required. The successful candidates will be responsible for teaching at one or more of the three levels of education (BSc, MSc, and PhD) as well as in the programme for continuing education. A good knowledge of both spoken and written French is required. If this is not the case, the applicant should be ready to learn French within a year. Fluency in English is essential and active or passive knowledge of other languages is an additional advantage.
For exact application procedure consult http://www.uclouvain.be/en-38120.html
Closing date : *11 January* 2008.

For more information, contact (ref: *SC/BIOL17/2008/842*) **Prof. R. Wesselingh (wesselingh@ecol.ucl.ac.be, +32 10 47 34 47),** (ref: *SC/BIOL13/2008/841*) **Prof. B. Knoops (knoops@bani.ucl.ac.be, +32 10 47 37 60), or for both references Prof. Th. Hance (hance@ecol.ucl.ac.be, +32 10 47 34 93).**

The Université catholique de Louvain is committed to equality of opportunity in employment.
W113959R

Get the latest jobs while playing with your BlackBerry®.

BlackBerry is a trademark of Research In Motion Limited.

naturejobs

nature publishing group npg

natu-oct-04-07-p000 29/9/07 7:44 Page 6

OPPORTUNITIES @ NIH THE NATIONAL INSTITUTES OF HEALTH

National Center for Research Resources

NATIONAL INSTITUTES OF HEALTH

Deputy Director, Clinical and Translational Research

THE POSITION: The National Center for Research Resources (NCRR) is seeking exceptional candidates for the position of Deputy Director, Clinical and Translational Research for the Center. The incumbent will lead NCRR efforts to integrate basic discoveries with clinical research and ensure that the resources supported by NCRR catalyze the advancement of biomedical research. He/She will advise the Director, NCRR, on the importance, policy implications, and program significance of current clinical and translational research issues, focusing on translation from basic research into pre-clinical studies and clinical trials, recommend changes in policy/operations, or follow-up actions. Areas of responsibility include sensitive biomedical and/or political issues that cut across the NIH. The NCRR provides laboratory scientists and clinical researchers with the environments and tools they need to understand, detect, treat, and prevent a wide range of diseases. **www.ncrr.nih.gov** This support enables discoveries that begin at a molecular and cellular level, move to animal-based studies, and then are translated to patient-oriented clinical research, resulting in cures and treatments for both common and rare diseases. This position offers a unique and exciting opportunity for an extremely capable individual to share responsibility in providing strong and visionary leadership to an organization dedicated to enhancing our understanding of health and disease, translating basic research into medical care, and improving human health. The Deputy Director, Clinical and Translational Research will be expected to represent the Director on a broad range of clinical and translational research issues related to the Center's activities before Members of Congress and their staffs, high level Government officials, leaders of national voluntary and professional health organizations, and leaders in business, science and academia.

QUALIFICATIONS REQUIRED: Applicants must possess an M.D., Ph.D., or equivalent degree, as well as senior-level research experience or knowledge of research programs moving research from the basic laboratory sciences into pre-clinical models and clinical trials. Candidates should be outstanding communicators and known and respected as distinguished individuals of outstanding competence. Applicants should also demonstrate the ability to think strategically, work collaboratively and use a consultative approach to problem solving and decision making.

SALARY/BENEFITS/OTHER INFORMATION: Salary is commensurate with experience and a full package of Civil Service benefits is available, including: retirement, health and life insurance, long term care insurance, leave and savings plan (401K equivalent). The National Institutes of Health inspires public confidence in science by maintaining high ethical principles. In addition to the Federal government's code of ethics, we have our own agency specific standards - check them out at the NIH Ethics web site. This position is subject to a background investigation.

HOW TO APPLY: A Curriculum Vitae, Bibliography, and two letters of recommendation must be received by **November 30, 2007**. Application packages should be sent to the **National Institutes of Health, National Center for Research Resources, ATTN: Bonnie Richards, 6701 Democracy Boulevard, Suite 1010, Bethesda, Maryland 20892**.

For further information, please call **(301) 435-0717**. All information provided by candidates will remain confidential and will not be released outside the NCRR search process without a signed release from candidates.

natu-oct-04-07-p000 29/9/07 7:45 Page 7

TENURE TRACK POSITION IN IMMUNOLOGY

The Experimental Immunology Branch (EIB), Center for Cancer Research (CCR) of the National Cancer Institute (NCI), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), Bethesda, MD, invites applicants to apply for a tenure track position in Immunology. The applicant should have a Ph.D., M.D. or equivalent degree, a strong record of scientific accomplishments, and the potential to establish an independent research program in any aspect of molecular or cellular immunology. The position provides salary and full funding to establish an independent research program, including laboratory space, equipment, budget, technical personnel, and support for fellows. Salary will be commensurate with education and experience. The EIB consists of 9 Principal Investigators: Alfred Singer, Triantafyllos Chavakis, Richard Hodes, Andre Nussenzweig, Paul Roche, David Segal, Stephen Shaw, Gene Shearer and Dinah Singer. Active research covers a wide range of areas of immunology including: thymic education and T cell differentiation, HIV-induced immunodeficiency, genetic recombination and chromosomal instability, inflammation biology, antigen presentation, receptor assembly and transport, signal transduction, and regulation of gene expression (http://rex.nci.nih.gov/RESEARCH/basic/eib/eibpage.htm). Scientific interactions are encouraged and occur extensively among members of EIB as well as with other scientists at the NIH. Applicants should send a CV and bibliography, outline of a proposed research program (no more than two pages), and three letters of recommendation to **Caroline McCabe, 10 Center Drive, Bldg. 10, Room 4B36, NIH, Bethesda, MD 20892-1360**. Applications must be received by **December 15, 2007**.

Tenure Track Investigator

The Lab of Receptor Biology and Gene Expression, Center for Cancer Research, National Cancer Institute, is recruiting for a position in the area of chromatin structure and function, chromosome biology, and nuclear architecture. The position is at the level of Tenure Track Principal Investigator, but senior investigators with expertise in the program area may be considered. Applicants should have a Ph.D. or M.D. degree, a strong publication record, and demonstrated potential in creative research. Salary will be commensurate with education and experience. Applications should be submitted as electronic files (Word or pdf docs), and should include a curriculum vitae, statement of research interests, and 3 letters of recommendation. Submit by Nov. 30, 2007 to: Christine Koch-Paiz, Lab of Receptor Biology & Gene Expression, National Cancer Institute, NIH, Bethesda MD 20892-5044, paizc@mail.nih.gov.

HIV and AIDS Malignancy Branch
Center for Cancer Research
Tenure Track or Tenure Eligible Position in Viral Oncogenesis

The HIV and AIDS Malignancy Branch (HAMB), NCI, is searching for a tenure track or tenure eligible investigator in the field of viral oncogenesis. It is anticipated that the investigator will establish an independent research program targeted to the study of viral-induced tumors, especially those associated with AIDS. The research program should be able to interface with the branch's existing clinical and basic programs in AIDS-associated malignancies. A particular interest will be for a research program in gammaherpesviruses, but other areas of viral oncogenesis will be considered as well. Current areas of laboratory research in HAMB focus on the molecular biology of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) and human papillomavirus (HPV), pathogenesis of tumors caused by these viruses, and the development of novel therapeutic interventions for HIV infection. The clinical research program in HAMB is primarily directed at AIDS-related malignancies. HAMB is located on the Bethesda campus of the NIH (http://ccr.cancer.gov/labs/lab.asp?labid=63). Candidates for the position should have an M.D./Ph.D., Ph.D., or M.D. and strong research credentials. Applicants for this position should submit a curriculum vitae including bibliography, a statement of research interests, a two-page outline of the proposed research program, and the names of three references to **Chairman, Search Committee, HAMB, NCI, Attention Jan Huque, 301-435-4627, fax 301-480-5955, Building 10, Rm. 6N106, 10 Center Drive, MSC 1868, Bethesda, MD 20892-1868 no later than November 18, 2007.** You may also e-mail your application to: **huquej@mail.nih.gov.**

NW113615R

OPPORTUNITIES @ NIH THE NATIONAL INSTITUTES OF HEALTH

ARGUS blinked

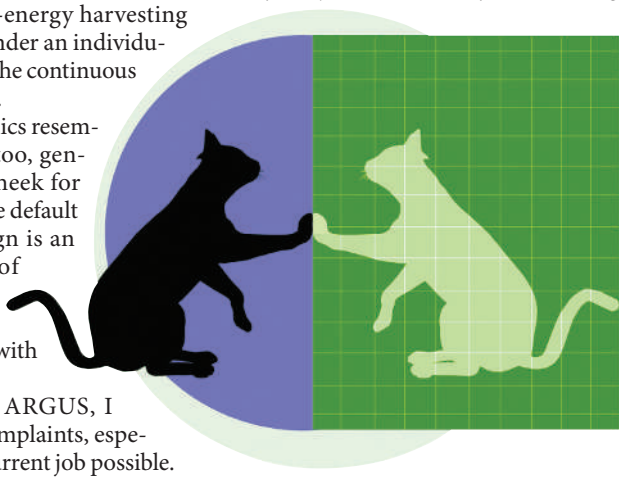
You're being watched.

Paul Di Filippo

My cat was watching me at my workstation. And so was everyone else in the world. Nowadays we all live in a real-time Panopticon. Thanks to ARGUS.

ARGUS is the ARchive of Globally Uploaded Sensoria, and it contains every second of what every person on Earth sees or hears, even while they're asleep. An array of deer-tick-sized cameras and mics, powered by ambient-energy harvesting and embedded just under an individual's skin, takes care of the continuous volitionless recording.

The cameras and mics resemble a small facial tattoo, generally one on each cheek for stereo processing. The default manufacturer's design is an iconographic Eye of Horus, but hardly anyone out of eight billion citizens sticks with the default.



JACEY

Growing up with ARGUS, I never had any real complaints, especially as it made my current job possible.

But then came that one disturbing day.

My name is Ross Strucker, and I'm an auteur. I turn the lives of ordinary people into art. Or I did, until I put down my digital toolkit forever.

The day ARGUS blinked, I was composing a romantic thriller. I was trying fruitlessly to find a shot in the ARGUS archives that included my two main players from a third perspective. That's often hard to do when only the two people in question are present together, regarding each other. Lots of times I can find surveillance-cam footage that does the trick. But not this time.

So, reluctantly, I turned to pet-cam footage. I generally dislike using footage from the Eyes of Horus installed in dogs, cats, pigeons and other animals, as it frequently represents weird camera angles and abrupt shifts in focus. But this time I found something suitable.

Satisfied yet tired, I took a break, and considered my palette of subsequent narrative choices. ARGUS offers so much to select from, after all.

The whole world in a gem. The many, many petabytes that make up ARGUS are mirrored across redundant sites, each store comprising 60 kilograms of artificial memory diamond, whose carbon-12/carbon-13 lattice is only half full after 50 years of global input.

The instant-by-instant wireless feed from an individual's Eyes of Horus, tagged with a unique civic identifier, flows steadily into ARGUS itself, becoming merged with the citizen's lifestream to date.

The overwhelming majority of ARGUS data are open-source: privacy and secrecy died as soon as ARGUS came online. Anything that one person knows or experiences can be known and used by anybody else.

My cat jumped into my lap, seeking attention I couldn't really spare. I was too busy pondering the fates of my characters, wondering how I could improve on the vast tapestry of raw realism contained in ARGUS.

The footage (we auteurs prefer the old-fashioned term) that every citizen provides is automatically tagged with a plethora of descriptive labels for every second, identifying its content a thousand different ways. Semantics-savvy retrieval engines can bring up selections effortlessly according to their commonplace content.

Show me what I had for dinner a year ago today.
What my ex-wife is doing right now.
Who met with the Emir of Paris at ten this morning.
When my son last took a bath.
What outfit Steffi Chubb is planning to wear to the Vatican Awards in Lagos tonight.

But my special auteurs' toolkit of semi-intelligent aesthetic agents allows me to select footage on a more arcane basis.

Show me a set of ironic responses to failed plans.
A set of nostalgic daydreamers in bucolic settings.
A set of locales that convey desuetude mixed with menace.
A set of stifled orgasms.

Out of the raw material trawled up from the depths of ARGUS and displayed on my wall-sized Coldfire monitor, I assemble narratives and stories. My work falls midway between the oneiric, surreal montages of such auteurs as The Culling House Collective, Armand Akimbo and the Voest Twins, and documentarians like Nilda Osborne, Focal Length Unlimited and the Informavore.

Just then, my cat decided it would get no affection from me, and chose instead to regard the ARGUS monitor with feline curiosity, looking at the screen as if it truly comprehended the cycling images from its animal compatriots on display there.

On a juvenile whim, I decided to create an endless-hall effect, the simple result of any camera trained on a live monitor accepting that camera's feed.

I was already in the pet-cam area of ARGUS, so it was simple to open a window onto my cat's lifestream.

But instead of the endless hall, I saw something impossible. On my screen appeared an image of my cat looking out of my monitor, as if my cat's onboard Eyes had been transmitting an image from a mirror.

What was ARGUS doing? What unknown glitch could possibly account for this?

And then it struck me: ARGUS was looking back at us. The digitized lifestreams inside the titanic archive had bootstrapped themselves into awareness. The simulacrum of the world had passed a tipping point of information density.

I grew dizzy, faint. I closed my eyes. When I opened them, the impossible cat looking intelligently out had been replaced by the endless hall I had expected.

Bored, my cat leaped down and the moving point-of-view on the monitor shifted accordingly.

I hurriedly shut off my system.

And I still haven't turned it back on. ■

With thanks to Charles Stross and Rudy Rucker for their seminal insights into lifelogs and lifeboxes.

Paul Di Filippo, a Rhode Island native, has lived in the lovecraftian stomping grounds of Providence for the past 31 years. He sold his first story in 1977, and well over 100 since. His new novel, *Cosmocopia*, will appear in January 2008.

©2007 Nature Publishing Group